

FIG. 1A

atg gct ttc tgc aca cag ttg atg ctt ctg ctg tgg aaa aat tac acc	48
Met Ala Phe Cys Thr Gln Leu Met Leu Leu Leu Trp Lys Asn Tyr Thr	
1 5 10 15	
tat cga cgg aga caa ccg atc caa cta cta gtg gag ttg ctt tgg ccc	96
Tyr Arg Arg Arg Gln Pro Ile Gln Leu Val Glu Leu Leu Trp Pro	
20 25 30	
ctc ttc ctc ttc ttc atc cta gtg gct gtc cgt cac tcc cac ccc cct	144
Leu Phe Leu Phe Phe Ile Leu Val Ala Val Arg His Ser His Pro Pro	
35 40 45	
ctg gag cat cac gaa tgc cac ttt cca aac aag cca tta cca tcg gcg	192
Leu Glu His His Glu Cys His Phe Pro Asn Lys Pro Leu Pro Ser Ala	
50 55 60	
ggc acg gtg ccc tgg ctg cag ggc ctt gtc tgc aac gta aac aac tcc	240
Gly Thr Val Pro Trp Leu Gln Gly Leu Val Cys Asn Val Asn Asn Ser	
65 70 75 80	
tgc ttc cag cac cca acg cct ggc gag aag cct ggg gtc ctg agt aac	288
Cys Phe Gln His Pro Thr Gly Glu Lys Pro Gly Val Leu Ser Asn	
85 90 95	
ttt aag gat tcc ttg atc tgc agg ctc ctc gct gat acc cgc aca gtg	336
Phe Lys Asp Ser Leu Ile Ser Arg Leu Leu Ala Asp Thr Arg Thr Val	
100 105 110	
ctc ggg ggc cac agc atc cag gac atg ctg gat gcc ctg ggg aaa ctg	384
Leu Gly Gly His Ser Ile Gln Asp Met Leu Asp Ala Leu Gly Lys Leu	
115 120 125	
atc ccc gtg ctc agg gca gtt gga ggt gga gca cga cca cag gag agt	432
Ile Pro Val Leu Arg Ala Val Gly Gly Gly Ala Arg Pro Gln Glu Ser	
130 135 140	
gac cag ccg acc agt caa ggg tca gtg act aag ctt ctg gag aag atc	480
Asp Gln Pro Thr Ser Gln Gly Ser Val Thr Lys Leu Leu Glu Lys Ile	
145 150 155 160	
ctg caa agg gca tcc ctg gat cct gtg ctg ggt caa gcc cag gat tct	528
Leu Gln Arg Ala Ser Leu Asp Pro Val Leu Gly Gln Ala Gln Asp Ser	
165 170 175	
atg aga aag ttc tca gat gct atc agg gat ctt gcc cag gag ctc ctg	576
Met Arg Lys Phe Ser Asp Ala Ile Arg Asp Leu Ala Gln Glu Leu Leu	
180 185 190	
aca ctg ccc agc ctg atg gag ctc cga gct ttg ctg cgg agg ccc cga	624
Thr Leu Pro Ser Leu Met Glu Leu Arg Ala Leu Leu Arg Arg Pro Arg	
195 200 205	

1000  
 900  
 800  
 700  
 600  
 500  
 400  
 300  
 200  
 100  
 0

1. The first part of the document is a list of names and their corresponding addresses. The names are listed in a column on the left, and the addresses are listed in a column on the right. The names are: John A. Smith, John B. Smith, John C. Smith, John D. Smith, John E. Smith, John F. Smith, John G. Smith, John H. Smith, John I. Smith, John J. Smith, John K. Smith, John L. Smith, John M. Smith, John N. Smith, John O. Smith, John P. Smith, John Q. Smith, John R. Smith, John S. Smith, John T. Smith, John U. Smith, John V. Smith, John W. Smith, John X. Smith, John Y. Smith, John Z. Smith. The addresses are: 123 Main St., 456 Main St., 789 Main St., 101 Main St., 202 Main St., 303 Main St., 404 Main St., 505 Main St., 606 Main St., 707 Main St., 808 Main St., 909 Main St., 1010 Main St., 1111 Main St., 1212 Main St., 1313 Main St., 1414 Main St., 1515 Main St., 1616 Main St., 1717 Main St., 1818 Main St., 1919 Main St., 2020 Main St., 2121 Main St., 2222 Main St., 2323 Main St., 2424 Main St., 2525 Main St., 2626 Main St., 2727 Main St., 2828 Main St., 2929 Main St., 3030 Main St., 3131 Main St., 3232 Main St., 3333 Main St., 3434 Main St., 3535 Main St., 3636 Main St., 3737 Main St., 3838 Main St., 3939 Main St., 4040 Main St., 4141 Main St., 4242 Main St., 4343 Main St., 4444 Main St., 4545 Main St., 4646 Main St., 4747 Main St., 4848 Main St., 4949 Main St., 5050 Main St., 5151 Main St., 5252 Main St., 5353 Main St., 5454 Main St., 5555 Main St., 5656 Main St., 5757 Main St., 5858 Main St., 5959 Main St., 6060 Main St., 6161 Main St., 6262 Main St., 6363 Main St., 6464 Main St., 6565 Main St., 6666 Main St., 6767 Main St., 6868 Main St., 6969 Main St., 7070 Main St., 7171 Main St., 7272 Main St., 7373 Main St., 7474 Main St., 7575 Main St., 7676 Main St., 7777 Main St., 7878 Main St., 7979 Main St., 8080 Main St., 8181 Main St., 8282 Main St., 8383 Main St., 8484 Main St., 8585 Main St., 8686 Main St., 8787 Main St., 8888 Main St., 8989 Main St., 9090 Main St., 9191 Main St., 9292 Main St., 9393 Main St., 9494 Main St., 9595 Main St., 9696 Main St., 9797 Main St., 9898 Main St., 9999 Main St.

ggg Gly	tca Ser	gct Ala	ggt Gly	tct Ser	ctg Leu	gag Glu	ctg Leu	gtt Val	tcg Ser	gag Glu	gcc Ala	ctc Leu	tgc Cys	agt Ser	acc Thr	672
210215220																
aag Lys	gga Gly	ccc Pro	agc Ser	agt Ser	cca Pro	ggg Gly	ggc Gly	ctg Leu	tcc Ser	ctc Leu	aat Asn	tgg Trp	tac Tyr	gaa Glu	gcc Ala	720
225230235240																
aac Asn	cag Gln	ctt Leu	aat Asn	gag Glu	ttc Phe	atg Met	ggg Gly	cca Pro	gag Glu	gtg Val	gcc Ala	cct Pro	gcc Ala	ctg Leu	cct Pro	768
245250255																
gac Asp	aac Asn	agt Ser	ctc Leu	agc Ser	cct Pro	gcc Ala	tgc Cys	tct Ser	gag Glu	ttt Phe	gtg Val	ggg Gly	aca Thr	ctg Leu	gat Asp	816
260265270																
gac Asp	cac His	cct Pro	gtg Val	tct Ser	cgg Arg	ctg Leu	ctc Leu	tgg Trp	agg Arg	cgc Arg	ctg Leu	aag Lys	cca Pro	ttg Leu	atc Ile	864
275280285																
ctc Leu	ggg Gly	aaa Lys	att Ile	ctc Leu	ttt Phe	gca Ala	cct Pro	gac Asp	aca Thr	aac Asn	ttc Phe	act Thr	cgg Arg	aag Lys	ctc Leu	912
290295300																
atg Met	gct Ala	cag Gln	gtg Val	aac Asn	cag Gln	acc Thr	ttc Phe	gag Glu	gag Glu	ctg Leu	gct Ala	ctg Leu	ttg Leu	agg Arg	gac Asp	960
305310315320																
cta Leu	cac His	gaa Glu	ctc Leu	tgg Trp	ggg Gly	gtg Val	ctg Leu	gga Gly	ccc Pro	cag Gln	atc Ile	ttc Phe	aac Asn	ttc Phe	atg Met	1008
325330335																
aat Asn	gac Asp	agt Ser	acc Thr	aac Asn	gtg Val	gcc Ala	atg Met	ctt Leu	cag Gln	agg Arg	ctt Leu	ctg Leu	gat Asp	gtg Val	ggg Gly	1056
340345350																
ggc Gly	aca Thr	ggg Gly	cag Gln	agg Arg	cag Gln	cag Gln	aca Thr	ccc Pro	aga Arg	gcc Ala	cag Gln	aag Lys	aag Lys	ttg Leu	gag Glu	1104
355360365																
gct Ala	atc Ile	aaa Lys	gac Asp	ttt Phe	ctg Leu	gat Asp	cct Pro	agt Ser	agg Arg	ggg Gly	ggc Gly	tac Tyr	agc Ser	tgg Trp	cgg Arg	1152
370375380																
gag Glu	gcc Ala	cac His	gca Ala	gat Asp	atg Met	gga Gly	cgc Arg	ctg Leu	gct Ala	gga Gly	atc Ile	cta Leu	gga Gly	caa Gln	atg Met	1200
385390395400																
atg Met	gag Glu	tgt Cys	gtg Val	tcc Ser	ctg Leu	gac Asp	aag Lys	ctg Leu	gag Glu	gct Ala	gtg Val	ccc Pro	tca Ser	gag Glu	gaa Glu	1248
405410415																

FIG. 1C

gct ctt gtg tcc cgt gcc ctg gag ctg ctg ggt gag cgc cgc ctc tgg	1296
Ala Leu Val Ser Arg Ala Leu Glu Leu Leu Gly Glu Arg Arg Leu Trp	
420 425 430	
gca ggc atc gtg ttc ctg agc cca gag cat cct ctg gac cca tcc gaa	1344
Ala Gly Ile Val Phe Leu Ser Pro Glu His Pro Leu Asp Pro Ser Glu	
435 440 445	
ctg tca tct cca gcc ctg agt cct ggc cac cta cga ttc aag att cga	1392
Leu Ser Ser Pro Ala Leu Ser Pro Gly His Leu Arg Phe Lys Ile Arg	
450 455 460	
atg gat atc gat gat gtc aca agg acc aat aag atc agg gac aag ttt	1440
Met Asp Ile Asp Asp Val Thr Arg Thr Asn Lys Ile Arg Asp Lys Phe	
465 470 475 480	
tgg gac cca ggt ccg tca gca gat cct ttc atg gac ctt cgg tat gtg	1488
Trp Asp Pro Gly Pro Ser Ala Asp Pro Phe Met Asp Leu Arg Tyr Val	
485 490 495	
tgg gga ggc ttc gtg tac ctg cag gac ctg ctg gag cag gca gct gtg	1536
Trp Gly Gly Phe Val Tyr Leu Gln Asp Leu Leu Glu Gln Ala Ala Val	
500 505 510	
cga gtg ctc ggt ggc ggg aac tcc cgc aca ggt ctc tac ctg cag cag	1584
Arg Val Leu Gly Gly Gly Asn Ser Arg Thr Gly Leu Tyr Leu Gln Gln	
515 520 525	
atg cca cac ccc tgc tac gtg gat gat gtg ttc ctg cgg gtg ctg agc	1632
Met Pro His Pro Cys Tyr Val Asp Asp Val Phe Leu Arg Val Leu Ser	
530 535 540	
cgg tct ctg cct ctg ttt ctg act ctg gcc tgg att tat tcg gtg gcg	1680
Arg Ser Leu Pro Leu Phe Leu Thr Leu Ala Trp Ile Tyr Ser Val Ala	
545 550 555 560	
ctc act gtg aag gcc gtg gtg cgt gag aaa gag aca cgg ctg cga gaa	1728
Leu Thr Val Lys Ala Val Val Arg Glu Lys Glu Thr Arg Leu Arg Glu	
565 570 575	
acc atg cgt gcg atg ggg ctg agc cgc gcg gtg ctc tgg ctt ggt tgg	1776
Thr Met Arg Ala Met Gly Leu Ser Arg Ala Val Leu Trp Leu Gly Trp	
580 585 590	
ttc ctc agc tgc ctg gga ccc ttc ctg gtc agc gct gcg ttg ctg gta	1824
Phe Leu Ser Cys Leu Gly Pro Phe Leu Val Ser Ala Ala Leu Leu Val	
595 600 605	
tta gtg ctt aag cta ggg aac atc ctt cct tac agc cac ccg gtt gta	1872
Leu Val Leu Lys Leu Gly Asn Ile Leu Pro Tyr Ser His Pro Val Val	
610 615 620	

FIG. 1D

atc ttc ctt ttc ttg gcg gcc ttc gcg gtg gcc acc gtc gct cag agt	1920
Ile Phe Leu Phe Leu Ala Ala Phe Ala Val Ala Thr Val Ala Ala Gln Ser	
625 630 635 640	
ttt ctg ctc agc gcc ttc ttc tcc agg gcc aat ctg gca gca gcc tgc	1968
Phe Leu Leu Ser Ala Phe Phe Ser Arg Ala Asn Leu Ala Ala Ala Cys	
645 650 655	
ggg ggg ctc gcc tat ttt gcg ctc tat ctg ccc tac gta ctg tgt gtg	2016
Gly Gly Leu Ala Tyr Phe Ala Leu Tyr Leu Pro Tyr Val Leu Cys Val	
660 665 670	
gcc tgg cgc gag cgc ctg cac ctg ggc gga ctc tta gct gcg agc ctg	2064
Ala Trp Arg Glu Arg Leu His Leu Gly Gly Leu Leu Ala Ala Ser Leu	
675 680 685	
ctg tcc cct gta gcc ttt ggc ttt gga tgc gaa agc ctg gcg cta cta	2112
Leu Ser Pro Val Ala Phe Gly Phe Gly Cys Glu Ser Leu Ala Leu Leu	
690 695 700	
gag gag cag gga gac ggg gct cag tgg cac aat ttg ggc aca ggc ccc	2160
Glu Glu Gln Gly Asp Gly Ala Gln Trp His Asn Leu Gly Thr Gly Pro	
705 710 715 720	
gcg gag gac gtc ttc agc ctg gcc cag gtg tct gcc ttc ctg ttg ctt	2208
Ala Glu Asp Val Phe Ser Leu Ala Gln Val Ser Ala Phe Leu Leu Leu	
725 730 735	
gat gcc gtc atc tac ggc ctt gcc ctc tgg tac cta gag gct gtg tgc	2256
Asp Ala Val Ile Tyr Gly Leu Ala Leu Trp Tyr Leu Glu Ala Val Cys	
740 745 750	
cca ggc cag tat gga atc cct gaa cca tgg aat ttc cct ttt cgg agg	2304
Pro Gly Gln Tyr Gly Ile Pro Glu Pro Trp Asn Phe Pro Phe Arg Arg	
755 760 765	
agc tac tgg tgt gga cct ggg cct ccc aag agt tct gtc ttg gcc cct	2352
Ser Tyr Trp Cys Gly Pro Gly Pro Pro Lys Ser Ser Val Leu Ala Pro	
770 775 780	
gcc cca caa gat ccc aag gtt ctg gtg gaa gag cca ccc ctt ggc ctg	2400
Ala Pro Gln Asp Pro Lys Val Leu Val Glu Glu Pro Pro Leu Gly Leu	
785 790 795 800	
gtt cct ggt gtc tcc att cga ggc ctg aag aaa cat ttt cgt ggc tgt	2448
Val Pro Gly Val Ser Ile Arg Gly Leu Lys Lys His Phe Arg Gly Cys	
805 810 815	
ccg cag cca gcc ctg caa gga ctc aac ctt gac ttc tac gaa ggc cac	2496
Pro Gln Pro Ala Leu Gln Gly Leu Asn Leu Asp Phe Tyr Glu Gly His	
820 825 830	

FIG. 1E

atc act gcc ttt ttg ggt cac aac ggg gct ggc aag aca acc aca ctg	2544
Ile Thr Ala Phe Leu Gly His Asn Gly Ala Gly Lys Thr Thr Thr Leu	
835 840 845	
tcc att ttg agt ggt ctc ttc cca ccc agt agt ggc tcg gcc tcc atc	2592
Ser Ile Leu Ser Gly Leu Phe Pro Pro Ser Ser Gly Ser Ala Ser Ile	
850 855 860	
ctg ggc cat gat gta caa acc aac atg gca gcc atc cgg ccc cac ctg	2640
Leu Gly His Asp Val Gln Thr Asn Met Ala Ala Ile Arg Pro His Leu	
865 870 875 880	
ggc atc tgc ccg cag tac aat gtg ctg ttt gat atg ctg aca gtg gaa	2688
Gly Ile Cys Pro Gln Tyr Asn Val Leu Phe Asp Met Leu Thr Val Glu	
885 890 895	
gaa cat gtt tgg ttc tat ggc cgt ttg aaa ggc gtg agt gca gcc gcc	2736
Glu His Val Trp Phe Tyr Gly Arg Leu Lys Gly Val Ser Ala Ala Ala	
900 905 910	
atg ggc ccc gag cgg gaa cgt ctg ata cgg gat gtg ggg ctt acc ctc	2784
Met Gly Pro Glu Arg Glu Arg Leu Ile Arg Asp Val Gly Leu Thr Leu	
915 920 925	
aag cgg gac aca cag aca cgc cac ctc tct ggt gga atg cag aga aaa	2832
Lys Arg Asp Thr Gln Thr Arg His Leu Ser Gly Gly Met Gln Arg Lys	
930 935 940	
ctt tct gtg gcc att gcc ttt gtg ggt ggc tct cgt gtg gtc atc atg	2880
Leu Ser Val Ala Ile Ala Phe Val Gly Gly Ser Arg Val Val Ile Met	
945 950 955 960	
gac gag ccc act gct ggt gtg gac ccc gct tcc cgc cgt ggc att tgg	2928
Asp Glu Pro Thr Ala Gly Val Asp Pro Ala Ser Arg Arg Gly Ile Trp	
965 970 975	
gaa ttg cta ctt aag tac aga gaa ggt cgg aca ctg att ctc tcc act	2976
Glu Leu Leu Leu Lys Tyr Arg Glu Gly Arg Thr Leu Ile Leu Ser Thr	
980 985 990	
cac cac ctg gat gag gca gag ctc ttg gga gat cgc gtg gcc atg gtg	3024
His His Leu Asp Glu Ala Glu Leu Leu Gly Asp Arg Val Ala Met Val	
995 1000 1005	
gca ggt ggc tct ttg tgc tgc tgt ggg tcc ccg ctt ttc ttg cgc cga	3072
Ala Gly Gly Ser Leu Cys Cys Cys Gly Ser Pro Leu Phe Leu Arg Arg	
1010 1015 1020	
cac ttg ggc tgc ggt tac tac ctg acc ctg gtg aag agt tct cag tcc	3120
His Leu Gly Cys Gly Tyr Tyr Leu Thr Leu Val Lys Ser Ser Gln Ser	
1025 1030 1035 1040	

2544 2592 2640 2688 2736 2784 2832 2880 2928 2976 3024 3072 3120



# FIG. 1G

ggc ctg ttt gca cag gtt gtg ttg cct gcc ctc ttt gtg ggc ctg gcc	3792
Gly Leu Phe Ala Gln Val Val Leu Pro Ala Leu Phe Val Gly Leu Ala	
1250 1255 1260	
ctg ttc ttc agc ctc att gtg cct cct ttt ggc cag tac cca ccc ctg	3840
Leu Phe Phe Ser Leu Ile Val Pro Pro Phe Gly Gln Tyr Pro Pro Leu	
1265 1270 1275 1280	
cag ctc agc cct gct atg tat ggc cct cag gtc tcg ttc ttc agt gag	3888
Gln Leu Ser Pro Ala Met Tyr Gly Pro Gln Val Ser Phe Phe Ser Glu	
1285 1290 1295	
gat gcc cct ggg gac ccc aac cgg atg aag ctg ctg gag gct ctg cta	3936
Asp Ala Pro Gly Asp Pro Asn Arg Met Lys Leu Leu Glu Ala Leu Leu	
1300 1305 1310	
ggg gag gct ggg ctg cag gaa ccc agt atg cag gac aaa gat gcc agg	3984
Gly Glu Ala Gly Leu Gln Glu Pro Ser Met Gln Asp Lys Asp Ala Arg	
1315 1320 1325	
gga tct gag tgt aca cac tcc cta gct tgc tac ttc acg gtc cct gag	4032
Gly Ser Glu Cys Thr His Ser Leu Ala Cys Tyr Phe Thr Val Pro Glu	
1330 1335 1340	
gtc cct cct gat gtg gcc agc atc ctg gcc agt ggc aac tgg acg cca	4080
Val Pro Pro Asp Val Ala Ser Ile Leu Ala Ser Gly Asn Trp Thr Pro	
1345 1350 1355 1360	
gaa tct cca tcc cca gct tgc caa tgc agt cag cct gga gcc cgc cgc	4128
Glu Ser Pro Ser Pro Ala Cys Gln Cys Ser Gln Pro Gly Ala Arg Arg	
1365 1370 1375	
ctg ttg cca gat tgc ccg gct gga gct ggg ggt cca cca ccc ccc cag	4176
Leu Leu Pro Asp Cys Pro Ala Gly Ala Gly Gly Pro Pro Pro Gln	
1380 1385 1390	
gct gtg gct ggc ttg ggg gag gtg gtc cag aac ctc act ggc cga aat	4224
Ala Val Ala Gly Leu Gly Glu Val Val Gln Asn Leu Thr Gly Arg Asn	
1395 1400 1405	
gtg tct gac ttt ttg gtg aag aca tac ccc agc ctg gtg cgc cga ggc	4272
Val Ser Asp Phe Leu Val Lys Thr Tyr Pro Ser Leu Val Arg Arg Gly	
1410 1415 1420	
cta aag acc aag aag tgg gtg gat gag gtc aga tat ggg ggc ttc tcc	4320
Leu Lys Thr Lys Lys Trp Val Asp Glu Val Arg Tyr Gly Gly Phe Ser	
1425 1430 1435 1440	
ctg gga ggc cga gat cca gac ctg ccc aca ggg cat gag gtg gtc cgc	4368
Leu Gly Gly Arg Asp Pro Asp Leu Pro Thr Gly His Glu Val Val Arg	
1445 1450 1455	

Figure 1 consists of 12 histograms arranged horizontally, labeled  $x_0$  through  $x_{11}$ . Each histogram shows the frequency of values for  $x_k$  ranging from 0 to 10. The distributions are approximately normal, centered at 5. The peak frequency (count) for each  $x_k$  is as follows:

$k$	Peak Count (at $x_k = 5$ )
0	10
1	10
2	10
3	10
4	10
5	10
6	10
7	10
8	10
9	10
10	10
11	12

aca	ttg	gca	gag	att	cgg	gca	ctg	ctg	agc	ccc	caa	cct	ggg	aat	gcg	4416	
Thr	Leu	Ala	Glu	Ile	Arg	Ala	Leu	Leu	Ser	Pro	Gln	Pro	Gly	Asn	Ala		
			1460						1465						1470		
cta	gac	cgt	atc	ctg	aac	aac	ctc	act	cag	tgg	gcc	ctt	ggc	ctt	gat	4464	
Leu	Asp	Arg	Ile	Leu	Asn	Asn	Leu	Thr	Gln	Trp	Ala	Leu	Gly	Leu	Asp		
			1475						1480						1485		
gct	cgg	aac	agc	ctc	aag	atc	tgg	ttc	aac	aac	aag	ggc	tgg	cat	gcc	4512	
Ala	Arg	Asn	Ser	Leu	Lys	Ile	Trp	Phe	Asn	Asn	Lys	Gly	Trp	His	Ala		
			1490						1495						1500		
atg	gtg	gcc	ttt	gtg	aac	cga	gcc	aac	aat	gga	ctc	cta	cat	gcc	ctc	4560	
Met	Val	Ala	Phe	Val	Asn	Arg	Ala	Asn	Asn	Gly	Leu	Leu	His	Ala	Leu		
1505						1510						1515			1520		
cta	cca	tct	ggt	cca	gtc	cgc	cat	gcc	cac	agc	atc	act	aca	ctc	aac	4608	
Leu	Pro	Ser	Gly	Pro	Val	Arg	His	Ala	His	Ser	Ile	Thr	Thr	Leu	Asn		
			1525						1530						1535		
cat	cct	ttg	aac	ttg	acc	aag	gag	cag	cta	tct	gaa	gct	aca	ctg	ata	4656	
His	Pro	Leu	Asn	Leu	Thr	Lys	Glu	Gln	Leu	Ser	Glu	Ala	Thr	Leu	Ile		
			1540						1545						1550		
gcc	tcc	tct	gtg	gat	gtc	ctt	gtc	tcc	atc	tgt	gtg	gtc	ttc	gcc	atg	4704	
Ala	Ser	Ser	Val	Asp	Val	Leu	Val	Ser	Ile	Cys	Val	Val	Phe	Ala	Met		
			1555						1560						1565		
tca	ttt	gtc	cca	gcc	agc	ttt	acc	ctg	gtc	ctc	ata	gag	gaa	cgc	atc	4752	
Ser	Phe	Val	Pro	Ala	Ser	Phe	Thr	Leu	Val	Leu	Ile	Glu	Glu	Arg	Ile		
			1570						1575						1580		
acc	aga	gcc	aag	cat	ctg	cag	ctg	gtc	agc	ggc	ctg	ccc	caa	acc	ctc	4800	
Thr	Arg	Ala	Lys	His	Leu	Gln	Leu	Val	Ser	Gly	Leu	Pro	Gln	Thr	Leu		
1585						1590						1595			1600		
tat	tgg	ctt	ggc	aac	ttc	ctc	tgg	gac	atg	tgt	aac	tac	ttg	gtg	gca	4848	
Tyr	Trp	Leu	Gly	Asn	Phe	Leu	Trp	Asp	Met	Cys	Asn	Tyr	Leu	Val	Ala		
			1605						1610						1615		
gtg	tgc	ata	gtg	gtg	ttc	atc	ttc	cta	gcc	ttt	cag	cag	aga	gcc	tat	4896	
Val	Cys	Ile	Val	Val	Phe	Ile	Phe	Leu	Ala	Phe	Gln	Gln	Arg	Ala	Tyr		
			1620						1625						1630		
gtg	gcc	cca	gag	aac	ctg	cct	gct	ctc	tta	ctc	ttg	ctt	ctg	ctg	tat	4944	
Val	Ala	Pro	Glu	Asn	Leu	Pro	Ala	Leu	Leu	Leu	Leu	Leu	Leu	Leu	Tyr		
			1635						1640						1645		
ggg	tgg	tct	atc	aca	cca	ctc	atg	tac	cca	gcc	tcc	ttc	ttc	ttc	tca	4992	
Gly	Trp	Ser	Ile	Thr	Pro	Leu	Met	Tyr	Pro	Ala	Ser	Phe	Phe	Phe	Ser		
1650						1655						1660					



Figure 1. The effect of the concentration of the *Agrobacterium* strain on the transformation efficiency of *Agrobacterium* strain 102. The *Agrobacterium* strain 102 was cultured in the YEA medium for 24 h at 28°C. The cell concentration was adjusted to 1.0 × 10<sup>8</sup> cells/ml. The cell suspension was mixed with the cell suspension of the *Agrobacterium* strain 102 at the concentration of 1.0 × 10<sup>8</sup> cells/ml. The mixture was then transformed into the *Agrobacterium* strain 102. The transformation efficiency was determined by the number of transformants per 10<sup>8</sup> cells. The results are shown in the figure. The error bars represent the standard deviation.

gtg	ccc	agc	acg	gcc	tat	gtg	gtg	ctc	acc	tgc	atc	aac	ctc	ttc	att	5040
Val	Pro	Ser	Thr	Ala	Tyr	Val	Val	Leu	Thr	Cys	Ile	Asn	Leu	Phe	Ile	
1665				1670				1675						1680		
ggc	atc	aat	agc	agc	atg	gcc	acc	ttc	gtg	cta	gaa	ctg	ctt	tca	gat	5088
Gly	Ile	Asn	Ser	Ser	Met	Ala	Thr	Phe	Val	Leu	Glu	Leu	Leu	Ser	Asp	
			1685					1690						1695		
cag	aac	ctg	caa	gaa	gtg	agc	cgg	atc	ctg	aaa	caa	gtg	ttt	ctt	att	5136
Gln	Asn	Leu	Gln	Glu	Val	Ser	Arg	Ile	Leu	Lys	Gln	Val	Phe	Leu	Ile	
		1700						1705					1710			
ttc	ccc	cac	ttt	tgc	ctt	ggc	cga	ggg	ctc	att	gac	atg	gtt	cgg	aac	5184
Phe	Pro	His	Phe	Cys	Leu	Gly	Arg	Gly	Leu	Ile	Asp	Met	Val	Arg	Asn	
	1715					1720					1725					
cag	gcc	atg	gca	gat	gcc	ttt	gag	cgc	tta	gga	gac	aag	caa	ttt	cag	5232
Gln	Ala	Met	Ala	Asp	Ala	Phe	Glu	Arg	Leu	Gly	Asp	Lys	Gln	Phe	Gln	
	1730				1735					1740						
tca	ccc	cta	cgc	tgg	gac	atc	att	ggc	aag	aac	ctc	ctg	gcc	atg	atg	5280
Ser	Pro	Leu	Arg	Trp	Asp	Ile	Ile	Gly	Lys	Asn	Leu	Leu	Ala	Met	Met	
1745				1750					1755					1760		
gcc	cag	gga	cct	ctg	ttc	ctc	ctc	atc	aca	ctc	ctg	ctc	caa	cac	cgc	5328
Ala	Gln	Gly	Pro	Leu	Phe	Leu	Leu	Ile	Thr	Leu	Leu	Leu	Gln	His	Arg	
		1765						1770					1775			
aac	cgt	ctc	ctg	cca	caa	tca	aaa	cca	aga	ctg	ctg	ccg	ccc	ctg	ggg	5376
Asn	Arg	Leu	Leu	Pro	Gln	Ser	Lys	Pro	Arg	Leu	Leu	Pro	Pro	Leu	Gly	
		1780					1785					1790				
gag	gag	gat	gag	gat	gtg	gct	caa	gag	cgt	gag	cgg	gtg	acc	aag	ggg	5424
Glu	Glu	Asp	Glu	Asp	Val	Ala	Gln	Glu	Arg	Glu	Arg	Val	Thr	Lys	Gly	
	1795					1800					1805					
gct	acc	cag	ggg	gat	gtg	cta	gtc	ctc	agg	gac	ttg	acc	aag	gtt	tac	5472
Ala	Thr	Gln	Gly	Asp	Val	Leu	Val	Leu	Arg	Asp	Leu	Thr	Lys	Val	Tyr	
	1810				1815					1820						
cgt	ggg	cag	agg	aac	cca	gct	gtg	gat	cgc	ctg	tgc	tta	ggg	atc	ccc	5520
Arg	Gly	Gln	Arg	Asn	Pro	Ala	Val	Asp	Arg	Leu	Cys	Leu	Gly	Ile	Pro	
1825				1830					1835					1840		
cct	ggg	gag	tgt	ttc	ggg	ctg	ctg	ggt	gtc	aac	ggg	gca	ggg	aag	aca	5568
Pro	Gly	Glu	Cys	Phe	Gly	Leu	Leu	Gly	Val	Asn	Gly	Ala	Gly	Lys	Thr	
		1845						1850					1855			
tcc	acc	ttc	cgc	atg	gtg	aca	ggg	gac	aca	ctg	ccc	agc	agt	ggg	gaa	5616
Ser	Thr	Phe	Arg	Met	Val	Thr	Gly</									

gca	gta	ctg	gca	ggc	cac	aac	gtg	gcc	cag	gag	cgg	tct	gcc	gca	cac	5664
Ala	Val	Leu	Ala	Gly	His	Asn	Val	Ala	Gln	Glu	Arg	Ser	Ala	Ala	His	
1875						1880						1885				
cgc	agc	atg	ggc	tac	tgt	ccc	cag	tct	gat	gcc	atc	ttc	gac	ctg	ctg	5712
Arg	Ser	Met	Gly	Tyr	Cys	Pro	Gln	Ser	Asp	Ala	Ile	Phe	Asp	Leu	Leu	
1890						1895						1900				
acc	ggc	cgg	gaa	cat	ctg	gaa	ctg	ttt	gct	cgc	ctg	cgc	ggg	gtg	ccc	5760
Thr	Gly	Arg	Glu	His	Leu	Glu	Leu	Phe	Ala	Arg	Leu	Arg	Gly	Val	Pro	
1905						1910						1915			1920	
gag	gcc	caa	gtt	gcc	cag	act	gcg	ctc	tct	ggc	ctg	gtg	cgc	ctg	ggc	5808
Glu	Ala	Gln	Val	Ala	Gln	Thr	Ala	Leu	Ser	Gly	Leu	Val	Arg	Leu	Gly	
			1925						1930						1935	
ctt	cct	agc	tat	gca	gac	cga	ccc	gcg	ggg	acc	tac	agc	gga	ggc	aac	5856
Leu	Pro	Ser	Tyr	Ala	Asp	Arg	Pro	Ala	Gly	Thr	Tyr	Ser	Gly	Gly	Asn	
			1940						1945						1950	
aaa	cgg	aag	ctg	gcg	aca	gcc	tta	gct	ctg	gtt	ggg	gac	cca	gct	gtg	5904
Lys	Arg	Lys	Leu	Ala	Thr	Ala	Leu	Ala	Leu	Val	Gly	Asp	Pro	Ala	Val	
1955						1960						1965				
gtc	ttt	ctg	gac	gag	ccc	acc	aca	ggc	atg	gac	cca	agt	gcg	cgg	cga	5952
Val	Phe	Leu	Asp	Glu	Pro	Thr	Thr	Gly	Met	Asp	Pro	Ser	Ala	Arg	Arg	
1970						1975						1980				
ttt	ctt	tgg	aac	agc	ttg	ctg	tcc	gtg	gtg	cgc	gag	ggc	cgc	tcc	gta	6000
Phe	Leu	Trp	Asn	Ser	Leu	Leu	Ser	Val	Val	Arg	Glu	Gly	Arg	Ser	Val	
1985						1990						1995			2000	
gtg	ctc	acg	tcg	cac	agc	atg	gag	gag	tgc	gaa	gcg	ctc	tgc	acg	cgc	6048
Val	Leu	Thr	Ser	His	Ser	Met	Glu	Glu	Cys	Glu	Ala	Leu	Cys	Thr	Arg	
			2005						2010						2015	
ctg	gcc	atc	atg	gtg	aac	ggg	cgg	ttc	cgc	tgt	ctg	gga	agc	tct	cag	6096
Leu	Ala	Ile	Met	Val	Asn	Gly	Arg	Phe	Arg	Cys	Leu	Gly	Ser	Ser	Gln	
			2020						2025						2030	
cat	ctc	aaa	ggc	agg	ttc	ggg	gct	ggc	cac	aca	ctg	act	cta	agg	gtc	6144
His	Leu	Lys	Gly	Arg	Phe	Gly	Ala	Gly	His	Thr	Leu	Thr	Leu	Arg	Val	
2035						2040						2045				
cca	ccg	gac	cag	cct	gag	ccg	gcg	ata	gcc	ttc	atc	agg	atc	aca	ttc	6192
Pro	Pro	Asp	Gln	Pro	Glu	Pro	Ala	Ile	Ala	Phe	Ile	Arg	Ile	Thr	Phe	
2050						2055						2060				
cct	ggg	gct	gaa	ctc	cgg	gag	gtg	cac	ggc	agc	cgt	ctg	cgc	ttc	caa	6240
Pro	Gly	Ala	Glu	Leu	Arg	Glu	Val	His	Gly	Ser	Arg	Leu	Arg	Phe	Gln	
2065						2070						2075			2080	

[illegible][illegible]

ctcagggg	cg	gcgcgctccc	tgccctgctgc	tgggcggagg	gaaggcggca	agagctgcgg	60
agcccc	tgga	agagcttcca	ggaacctgc	gctgtgggat	aaaggaatga	ggttcagaaa	120
ggggcaggga	gttgcccc	gca	gccgcaccgc	acgtcttcag	cccgaccgtt	gtcctgacct	180
ctctgtccc	g	tccccgc	ccc	agtctcacc	atg gcc ttc tgg aca cag ctg atg		233
					Met Ala Phe Trp Thr Gln Leu Met		
					1	5	
ctg ctg ctc tgg aag aat ttc atg tat cgc cgg aga cag ccg gtc cag	281						
Leu Leu Leu Trp Lys Asn Phe Met Tyr Arg Arg Arg Gln Pro Val Gln							
10	15	20					
ctc ctg gtc gaa ttg ctg tgg cct ctc ttc ctc ttc ttc atc ctg gtg	329						
Leu Leu Val Glu Leu Leu Trp Pro Leu Phe Leu Phe Phe Ile Leu Val							
25	30	35	40				
gct gtt cgc cac tcc cac ccg ccc ctg gag cac cat gaa tgc cac ttc	377						
Ala Val Arg His Ser His Pro Pro Leu Glu His His Glu Cys His Phe							
45	50	55					
cca aac aag cca ctg cca tcg gcg ggc acc gtg ccc tgg ctc cag ggt	425						
Pro Asn Lys Pro Leu Pro Ser Ala Gly Thr Val Pro Trp Leu Gln Gly							
60	65	70					
ctc atc tgt aat gtg aac aac acc tgc ttt ccg cag ctg aca ccg ggc	473						
Leu Ile Cys Asn Val Asn Asn Thr Cys Phe Pro Gln Leu Thr Pro Gly							
75	80	85					
gag gag ccc ggg cgc ctg agc aac ttc aac gac tcc ctg gtc tcc cgg	521						
Glu Glu Pro Gly Arg Leu Ser Asn Phe Asn Asp Ser Leu Val Ser Arg							
90	95	100					
ctg cta gcc gat gcc cgc act gtg ctg gga ggg gcc agt gcc cac agg	569						
Leu Leu Ala Asp Ala Arg Thr Val Leu Gly Gly Ala Ser Ala His Arg							
105	110	115	120				
acg ctg gct ggc cta ggg aag ctg atc gcc acg ctg agg gct gca cgc	617						
Thr Leu Ala Gly Leu Gly Lys Leu Ile Ala Thr Leu Arg Ala Ala Arg							
125	130	135					
agc acg gcc cag cct caa cca acc aag cag tct cca ctg gaa cca ccc	665						
Ser Thr Ala Gln Pro Gln Pro Thr Lys Gln Ser Pro Leu Glu Pro Pro							
140	145	150					
atg ctg gat gtc gcg gag ctg ctg acg tca ctg ctg cgc acg gaa tcc	713						
Met Leu Asp Val Ala Glu Leu Leu Thr Ser Leu Leu Arg Thr Glu Ser							
155	160	165					
ctg ggg ttg gca ctg ggc caa gcc cag gag ccc ttg cac agc ttg ttg	761						
Leu Gly Leu Ala Leu Gly Gln Ala Gln Glu Pro Leu His Ser Leu Leu							
170	175	180					

[illegible]

gag Glu 185	gcc Ala	gct Ala	gag Glu	gac Asp	ctg Leu 190	gcc Ala	cag Gln	gag Glu	ctc Leu	ctc Leu 195	gcg Ala	ctg Leu	cgc Arg	agc Ser	ctg Leu 200	809
gtg Val	gag Glu	ctt Leu	cgg Arg	gca Ala 205	ctg Leu	ctg Leu	cag Gln	aga Arg	ccc Pro 210	cga Arg	ggg Gly	acc Thr	agc Ser	ggc Gly 215	ccc Pro	857
ctg Leu	gag Glu	ttg Leu	ctg Leu 220	tca Ser	gag Glu	gcc Ala	ctc Leu	tgc Cys 225	agt Ser	gtc Val	agg Arg	gga Gly	cct Pro 230	agc Ser	agc Ser	905
aca Thr	gtg Val	ggc Gly 235	ccc Pro	tcc Ser	ctc Leu	aac Asn	tgg Trp 240	tac Tyr	gag Glu	gct Ala	agt Ser	gac Asp 245	ctg Leu	atg Met	gag Glu	953
ctg Leu 250	gtg Val	ggg Gly	cag Gln	gag Glu	cca Pro	gaa Glu 255	tcc Ser	gcc Ala	ctg Leu	cca Pro	gac Asp 260	agc Ser	agc Ser	ctg Leu	agc Ser	1001
ccc Pro 265	gcc Ala	tgc Cys	tcg Ser	gag Glu 270	ctg Leu	att Ile	gga Gly	gcc Ala	ctg Leu	gac Asp 275	agc Ser	cac His	ccg Pro	ctg Leu	tcc Ser 280	1049
cgc Arg	ctg Leu	ctc Leu	tgg Trp 285	aga Arg	cgc Arg	ctg Leu	aag Lys	cct Pro	ctg Leu 290	atc Ile	ctc Leu	ggg Gly	aag Lys	cta Leu 295	ctc Leu	1097
ttt Phe	gca Ala	cca Pro	gat Asp 300	aca Thr	cct Pro	ttt Phe	acc Thr	cgg Arg 305	aag Lys	ctc Leu	atg Met	gcc Ala	cag Gln 310	gtg Val	aac Asn	1145
cgg Arg	acc Thr 315	ttc Phe	gag Glu	gag Glu	ctc Leu	acc Thr	ctg Leu 320	ctg Leu	agg Arg	gat Asp	gtc Val	cgg Arg 325	gag Glu	gtg Val	tgg Trp	1193
gag Glu 330	atg Met	ctg Leu	gga Gly	ccc Pro	cgg Arg	atc Ile 335	ttc Phe	acc Thr	ttc Phe	atg Met	aac Asn 340	gac Asp	agt Ser	tcc Ser	aat Asn	1241
gtg Val 345	gcc Ala	atg Met	ctg Leu	cag Gln	cgg Arg 350	ctc Leu	ctg Leu	cag Gln	atg Met	cag Gln 355	gat Asp	gaa Glu	gga Gly	aga Arg	agg Arg 360	1289
cag Gln	ccc Pro	aga Arg	cct Pro	gga Gly 365	ggc Gly	cgg Arg	gac Asp	cac His	atg Met 370	gag Glu	gcc Ala	ctg Leu	cga Arg	tcc Ser 375	ttt Phe	1337
ctg Leu	gac Asp	cct Pro	ggg Gly 380	agc Ser	ggt Gly	ggc Gly	tac Tyr	agc Ser 385	tgg Trp	cag Gln	gac Asp	gca Ala	cac His 390	gct Ala	gat Asp	1385

[illegible]

gtg	ggg	cac	ctg	gtg	ggc	acg	ctg	ggc	cga	gtg	acg	gag	tgc	ctg	tcc	1433
Val	Gly	His	Leu	Val	Gly	Thr	Leu	Gly	Arg	Val	Thr	Glu	Cys	Leu	Ser	
		395				400						405				
ttg	gac	aag	ctg	gag	gcg	gca	ccc	tca	gag	gca	gcc	ctg	gtg	tcg	cgg	1481
Leu	Asp	Lys	Leu	Glu	Ala	Ala	Pro	Ser	Glu	Ala	Ala	Leu	Val	Ser	Arg	
		410				415						420				
gcc	ctg	caa	ctg	ctc	gcg	gaa	cat	cga	ttc	tgg	gcc	ggc	gtc	gtc	ttc	1529
Ala	Leu	Gln	Leu	Leu	Ala	Glu	His	Arg	Phe	Trp	Ala	Gly	Val	Val	Phe	
		425				430						435		440		
ttg	gga	cct	gag	gac	tct	tca	gac	ccc	aca	gag	cac	cca	acc	cca	gac	1577
Leu	Gly	Pro	Glu	Asp	Ser	Ser	Asp	Pro	Thr	Glu	His	Pro	Thr	Pro	Asp	
				445						450				455		
ctg	ggc	ccc	ggc	cac	gtg	cgc	atc	aaa	atc	cgc	atg	gac	att	gac	gtg	1625
Leu	Gly	Pro	Gly	His	Val	Arg	Ile	Lys	Ile	Arg	Met	Asp	Ile	Asp	Val	
		460						465						470		
gtc	acg	agg	acc	aat	aag	atc	agg	gac	agg	ttt	tgg	gac	cct	ggc	cca	1673
Val	Thr	Arg	Thr	Asn	Lys	Ile	Arg	Asp	Arg	Phe	Trp	Asp	Pro	Gly	Pro	
		475				480						485				
gcc	gcg	gac	ccc	ctg	acc	gac	ctg	cgc	tac	gtg	tgg	ggc	ggc	ttc	gtg	1721
Ala	Ala	Asp	Pro	Leu	Thr	Asp	Leu	Arg	Tyr	Val	Trp	Gly	Gly	Phe	Val	
		490				495						500				
tac	ctg	caa	gac	ctg	gtg	gag	cgt	gca	gcc	gtc	cgc	gtg	ctc	agc	ggc	1769
Tyr	Leu	Gln	Asp	Leu	Val	Glu	Arg	Ala	Ala	Val	Arg	Val	Leu	Ser	Gly	
		505				510						515		520		
gcc	aac	ccc	cgg	gcc	ggc	ctc	tac	ctg	cag	cag	atg	ccc	tat	ccg	tgc	1817
Ala	Asn	Pro	Arg	Ala	Gly	Leu	Tyr	Leu	Gln	Gln	Met	Pro	Tyr	Pro	Cys	
				525						530				535		
tat	gtg	gac	gac	gtg	ttc	ctg	cgt	gtg	ctg	agc	cgg	tcg	ctg	ccg	ctc	1865
Tyr	Val	Asp	Asp	Val	Phe	Leu	Arg	Val	Leu	Ser	Arg	Ser	Leu	Pro	Leu	
		540						545						550		
ttc	ctg	acg	ctg	gcc	tgg	atc	tac	tcc	gtg	aca	ctg	aca	gtg	aag	gcc	1913
Phe	Leu	Thr	Leu	Ala	Trp	Ile	Tyr	Ser	Val	Thr	Leu	Thr	Val	Lys	Ala	
		555				560						565				
gtg	gtg	cgg	gag	aag	gag	acg	cgg	ctg	cgg	gac	acc	atg	cgc	gcc	atg	1961
Val	Val	Arg	Glu	Lys	Glu	Thr	Arg	Leu	Arg	Asp	Thr	Met	Arg	Ala	Met	
		570				575						580				
ggg	ctc	agc	cgc	gcg	gtg	ctc	tgg	cta	ggc	tgg	ttc	ctc	agc	tgc	ctc	2009
Gly	Leu	Ser	Arg	Ala	Val	Leu	Trp	Leu	Gly	Trp	Phe	Leu	Ser	Cys	Leu	
		585				590						595		600		

FIG. 2D

ggg ccc ttc ctg ctc agc gcc gcg ctg ctg gtt ctg gtg ctc aag ctg	2057
Gly Pro Phe Leu Leu Ser Ala Ala Leu Leu Val Leu Val Leu Lys Leu	
605 610 615	
ggg gac atc ctc ccc tac agc cac ccg ggc gtg gtc ttc ctg ttc ttg	2105
Gly Asp Ile Leu Pro Tyr Ser His Pro Gly Val Val Phe Leu Phe Leu	
620 625 630	
gca gcc ttc gcg gtg gcc acg gtg acc cag agc ttc ctg ctc agc gcc	2153
Ala Ala Phe Ala Val Ala Thr Val Thr Gln Ser Phe Leu Leu Ser Ala	
635 640 645	
ttc ttc tcc cgc gcc aac ctg gct gcg gcc tgc ggc ggc ctg gcc tac	2201
Phe Phe Ser Arg Ala Asn Leu Ala Ala Ala Cys Gly Gly Leu Ala Tyr	
650 655 660	
ttc tcc ctc tac ctg ccc tac gtg ctg tgt gtg gct tgg cgg gac cgg	2249
Phe Ser Leu Tyr Leu Pro Tyr Val Leu Cys Val Ala Trp Arg Asp Arg	
665 670 675 680	
ctg ccc gcg ggt ggc cgc gtg gcc gcg agc ctg ctg tcg ccc gtg gcc	2297
Leu Pro Ala Gly Gly Arg Val Ala Ala Ser Leu Leu Ser Pro Val Ala	
685 690 695	
ttc ggc ttc ggc tgc gag agc ctg gct ctg ctg gag gag cag ggc gag	2345
Phe Gly Phe Gly Cys Glu Ser Leu Ala Leu Leu Glu Glu Gln Gly Glu	
700 705 710	
ggc gcg cag tgg cac aac gtg ggc acc cgg cct acg gca gac gtc ttc	2393
Gly Ala Gln Trp His Asn Val Gly Thr Arg Pro Thr Ala Asp Val Phe	
715 720 725	
agc ctg gcc cag gtc tct ggc ctt ctg ctg ctg gac gcg gcg ctc tac	2441
Ser Leu Ala Gln Val Ser Gly Leu Leu Leu Leu Asp Ala Ala Leu Tyr	
730 735 740	
ggc ctc gcc acc tgg tac ctg gaa gct gtg tgc cca ggc cag tac ggg	2489
Gly Leu Ala Thr Trp Tyr Leu Glu Ala Val Cys Pro Gly Gln Tyr Gly	
745 750 755 760	
atc cct gaa cca tgg aat ttt cct ttt cgg agg agc tac tgg tgc gga	2537
Ile Pro Glu Pro Trp Asn Phe Pro Phe Arg Arg Ser Tyr Trp Cys Gly	
765 770 775	
cct cgg ccc ccc aag agt cca gcc cct tgc ccc acc ccg ctg gac cca	2585
Pro Arg Pro Pro Lys Ser Pro Ala Pro Cys Pro Thr Pro Leu Asp Pro	
780 785 790	
aag gtg ctg gta gaa gag gca ccg ccc ggc ctg agt cct ggc gta tcc	2633
Lys Val Leu Val Glu Glu Ala Pro Pro Gly Leu Ser Pro Gly Val Ser	
795 800 805	

009541-1299

FIG. 2E

gtt cgc agc ctg gag aag cgc ttt cct gga agc ccg cag cca gcc ctg	2681
Val Arg Ser Leu Glu Lys Arg Phe Pro Gly Ser Pro Gln Pro Ala Leu	
810 815 820	
cgg ggg ctc agc ctg gac ttc tac cag ggc cac atc acc gcc ttc ctg	2729
Arg Gly Leu Ser Leu Asp Phe Tyr Gln Gly His Ile Thr Ala Phe Leu	
825 830 835 840	
ggc cac aac ggg gcc ggc aag acc acc acc ctg tcc atc ttg agt ggc	2777
Gly His Asn Gly Ala Gly Lys Thr Thr Thr Leu Ser Ile Leu Ser Gly	
845 850 855	
ctc ttc cca ccc agt ggt ggc tct gcc ttc atc ctg ggc cac gac gtc	2825
Leu Phe Pro Pro Ser Gly Gly Ser Ala Phe Ile Leu Gly His Asp Val	
860 865 870	
cgc tcc agc atg gcc gcc atc cgg ccc cac ctg ggc gtc tgt cct cag	2873
Arg Ser Ser Met Ala Ala Ile Arg Pro His Leu Gly Val Cys Pro Gln	
875 880 885	
tac aac gtg ctg ttt gac atg ctg acc gtg gac gag cac gtc tgg ttc	2921
Tyr Asn Val Leu Phe Asp Met Leu Thr Val Asp Glu His Val Trp Phe	
890 895 900	
tat ggg cgg ctg aag ggt ctg agt gcc gct gta gtg ggc ccc gag cag	2969
Tyr Gly Arg Leu Lys Gly Leu Ser Ala Ala Val Val Gly Pro Glu Gln	
905 910 915 920	
gac cgt ctg ctg cag gat gtg ggg ctg gtc tcc aag cag agt gtg cag	3017
Asp Arg Leu Leu Gln Asp Val Gly Leu Val Ser Lys Gln Ser Val Gln	
925 930 935	
act cgc cac ctc tct ggt ggg atg caa cgg aag ctg tcc gtg gcc att	3065
Thr Arg His Leu Ser Gly Gly Met Gln Arg Lys Leu Ser Val Ala Ile	
940 945 950	
gcc ttt gtg ggc ggc tcc caa gtt gtt atc ctg gac gag cct acg gct	3113
Ala Phe Val Gly Gly Ser Gln Val Val Ile Leu Asp Glu Pro Thr Ala	
955 960 965	
ggc gtg gat cct gct tcc cgc cgc ggt att tgg gag ctg ctg ctc aaa	3161
Gly Val Asp Pro Ala Ser Arg Arg Gly Ile Trp Glu Leu Leu Leu Lys	
970 975 980	
tac cga gaa ggt cgc acg ctg atc ctc tcc acc cac cac ctg gat gag	3209
Tyr Arg Glu Gly Arg Thr Leu Ile Leu Ser Thr His His Leu Asp Glu	
985 990 995 1000	
gca gag ctg ctg gga gac cgt gtg gct gtg gtg gca ggt ggc cgc ttg	3257
Ala Glu Leu Leu Gly Asp Arg Val Ala Val Val Ala Gly Gly Arg Leu	
1005 1010 1015	

1000  
 900  
 800  
 700  
 600  
 500  
 400  
 300  
 200  
 100  
 0



[illegible]

tgc	tgc	tgt	ggc	tcc	cca	ctc	ttc	ctg	cgc	cgt	cac	ctg	ggc	tcc	ggc	3305
Cys	Cys	Cys	Gly	Ser	Pro	Leu	Phe	Leu	Arg	Arg	His	Leu	Gly	Ser	Gly	
			1020					1025					1030			
tac	tac	ctg	acg	ctg	gtg	aag	gcc	cgc	ctg	ccc	ctg	acc	acc	aat	gag	3353
Tyr	Tyr	Leu	Thr	Leu	Val	Lys	Ala	Arg	Leu	Pro	Leu	Thr	Thr	Asn	Glu	
		1035					1040					1045				
aag	gct	gac	act	gac	atg	gag	ggc	agt	gtg	gac	acc	agg	cag	gaa	aag	3401
Lys	Ala	Asp	Thr	Asp	Met	Glu	Gly	Ser	Val	Asp	Thr	Arg	Gln	Glu	Lys	
	1050					1055				1060						
aag	aat	ggc	agc	cag	ggc	agc	aga	gtc	ggc	act	cct	cag	ctg	ctg	gcc	3449
Lys	Asn	Gly	Ser	Gln	Gly	Ser	Arg	Val	Gly	Thr	Pro	Gln	Leu	Leu	Ala	
1065				1070					1075						1080	
ctg	gta	cag	cac	tgg	gtg	ccc	ggg	gca	cgg	ctg	gtg	gag	gag	ctg	cca	3497
Leu	Val	Gln	His	Trp	Val	Pro	Gly	Ala	Arg	Leu	Val	Glu	Glu	Leu	Pro	
			1085					1090						1095		
cac	gag	ctg	gtg	ctg	gtg	ctg	ccc	tac	acg	ggc	gcc	cat	gac	ggc	agc	3545
His	Glu	Leu	Val	Leu	Val	Leu	Pro	Tyr	Thr	Gly	Ala	His	Asp	Gly	Ser	
		1100						1105					1110			
ttc	gcc	aca	ctc	ttc	cga	gag	cta	gac	acg	cgg	ctg	gcg	gag	ctg	agg	3593
Phe	Ala	Thr	Leu	Phe	Arg	Glu	Leu	Asp	Thr	Arg	Leu	Ala	Glu	Leu	Arg	
	1115					1120					1125					
ctc	act	ggc	tac	ggg	atc	tcc	gac	acc	agc	ctc	gag	gag	atc	ttc	ctg	3641
Leu	Thr	Gly	Tyr	Gly	Ile	Ser	Asp	Thr	Ser	Leu	Glu	Glu	Ile	Phe	Leu	
	1130				1135					1140						
aag	gtg	gtg	gag	gag	tgt	gct	gcg	gac	aca	gat	atg	gag	gat	ggc	agc	3689
Lys	Val	Val	Glu	Glu	Cys	Ala	Ala	Asp	Thr	Asp	Met	Glu	Asp	Gly	Ser	
1145				1150				1155						1160		
tgc	ggg	cag	cac	cta	tgc	aca	ggc	att	gct	ggc	cta	gac	gta	acc	ctg	3737
Cys	Gly	Gln	His	Leu	Cys	Thr	Gly	Ile	Ala	Gly	Leu	Asp	Val	Thr	Leu	
			1165					1170					1175			
cgg	ctc	aag	atg	cgc	cca	cag	gag	aca	gcg	ctg	gag	aac	ggg	gaa	cca	3785
Arg	Leu	Lys	Met	Pro	Pro	Gln	Glu	Thr	Ala	Leu	Glu	Asn	Gly	Glu	Pro	
		1180					1185					1190				
gct	ggg	tca	gcc	cca	gag	act	gac	cag	ggc	tct	ggg	cca	gac	gcc	gtg	3833
Ala	Gly	Ser	Ala	Pro	Glu	Thr	Asp	Gln	Gly	Ser	Gly	Pro	Asp	Ala	Val	
	1195				1200						1205					
ggc	cgg	gta	cag	ggc	tgg	gca	ctg	acc	cgc	cag	cag	ctc	cag	gcc	ctg	3881
Gly	Arg	Val	Gln	Gly	Trp	Ala										

項目	1990年	1991年	1992年	1993年	1994年	1995年	1996年	1997年	1998年	1999年	2000年	2001年	2002年	2003年	2004年	2005年	2006年	2007年	2008年	2009年	2010年	2011年	2012年	2013年	2014年	2015年	2016年	2017年	2018年	2019年	2020年	2021年	2022年	2023年	2024年	2025年	2026年	2027年	2028年	2029年	2030年	2031年	2032年	2033年	2034年	2035年	2036年	2037年	2038年	2039年	2040年	2041年	2042年	2043年	2044年	2045年	2046年	2047年	2048年	2049年	2050年	2051年	2052年	2053年	2054年	2055年	2056年	2057年	2058年	2059年	2060年	2061年	2062年	2063年	2064年	2065年	2066年	2067年	2068年	2069年	2070年	2071年	2072年	2073年	2074年	2075年	2076年	2077年	2078年	2079年	2080年	2081年	2082年	2083年	2084年	2085年	2086年	2087年	2088年	2089年	2090年	2091年	2092年	2093年	2094年	2095年	2096年	2097年	2098年	2099年	2100年																																																								
人口	120,000,000	121,000,000	122,000,000	123,000,000	124,000,000	125,000,000	126,000,000	127,000,000	128,000,000	129,000,000	130,000,000	131,000,000	132,000,000	133,000,000	134,000,000	135,000,000	136,000,000	137,000,000	138,000,000	139,000,000	140,000,000	141,000,000	142,000,000	143,000,000	144,000,000	145,000,000	146,000,000	147,000,000	148,000,000	149,000,000	150,000,000	151,000,000	152,000,000	153,000,000	154,000,000	155,000,000	156,000,000	157,000,000	158,000,000	159,000,000	160,000,000	161,000,000	162,000,000	163,000,000	164,000,000	165,000,000	166,000,000	167,000,000	168,000,000	169,000,000	170,000,000	171,000,000	172,000,000	173,000,000	174,000,000	175,000,000	176,000,000	177,000,000	178,000,000	179,000,000	180,000,000	181,000,000	182,000,000	183,000,000	184,000,000	185,000,000	186,000,000	187,000,000	188,000,000	189,000,000	190,000,000	191,000,000	192,000,000	193,000,000	194,000,000	195,000,000	196,000,000	197,000,000	198,000,000	199,000,000	200,000,000	201,000,000	202,000,000	203,000,000	204,000,000	205,000,000	206,000,000	207,000,000	208,000,000	209,000,000	210,000,000	211,000,000	212,000,000	213,000,000	214,000,000	215,000,000	216,000,000	217,000,000	218,000,000	219,000,000	220,000,000	221,000,000	222,000,000	223,000,000	224,000,000	225,000,000	226,000,000	227,000,000	228,000,000	229,000,000	230,000,000	231,000,000	232,000,000	233,000,000	234,000,000	235,000,000	236,000,000	237,000,000	238,000,000	239,000,000	240,000,000	241,000,000	242,000,000	243,000,000	244,000,000	245,000,000	246,000,000	247,000,000	248,000,000	249,000,000	250,000,000	251,000,000	252,000,000	253,000,000	254,000,000	255,000,000	256,000,000	257,000,000	258,000,000	259,000,000	260,000,000	261,000,000	262,000,000	263,000,000	264,000,000	265,000,000	266,000,000	267,000,000	268,000,000	269,000,000	270,000,000	271,000,000	272,000,000	273,000,000	274,000,000	275,000,000	276,000,000	277,000,000	278,000,000	279,000,000	280,000,000	281,000,000	282,000,000	283,000,000	284,000,000	285,000,000	286,000,000

ctt	ctc	aag	cgc	ttt	ctg	ctt	gcc	cgc	cgc	agc	cgc	cgc	ggc	ctg	ttc	3929
Leu	Leu	Lys	Arg	Phe	Leu	Leu	Ala	Arg	Arg	Ser	Arg	Arg	Gly	Leu	Phe	
1225					1230					1235					1240	
gcc	cag	atc	gtg	ctg	cct	gcc	ctc	ttt	gtg	ggc	ctg	gcc	ctc	gtg	ttc	3977
Ala	Gln	Ile	Val	Leu	Pro	Ala	Leu	Phe	Val	Gly	Leu	Ala	Leu	Val	Phe	
				1245					1250					1255		
agc	ctc	atc	gtg	cct	cct	ttc	ggg	cac	tac	ccg	gct	ctg	cgg	ctc	agt	4025
Ser	Leu	Ile	Val	Pro	Pro	Phe	Gly	His	Tyr	Pro	Ala	Leu	Arg	Leu	Ser	
			1260					1265					1270			
ccc	acc	atg	tac	ggt	gct	cag	gtg	tcc	ttc	ttc	agt	gag	gac	gcc	cca	4073
Pro	Thr	Met	Tyr	Gly	Ala	Gln	Val	Ser	Phe	Phe	Ser	Glu	Asp	Ala	Pro	
			1275				1280					1285				
ggg	gac	cct	gga	cgt	gcc	cgg	ctg	ctc	gag	gcg	ctg	ctg	cag	gag	gca	4121
Gly	Asp	Pro	Gly	Arg	Ala	Arg	Leu	Leu	Glu	Ala	Leu	Leu	Gln	Glu	Ala	
					1290		1295				1300					
gga	ctg	gag	gag	ccc	cca	gtg	cag	cat	agc	tcc	cac	agg	ttc	tcg	gca	4169
Gly	Leu	Glu	Glu	Pro	Pro	Val	Gln	His	Ser	Ser	His	Arg	Phe	Ser	Ala	
1305					1310					1315					1320	
cca	gaa	gtt	cct	gct	gaa	gtg	gcc	aag	gtc	ttg	gcc	agt	ggc	aac	tgg	4217
Pro	Glu	Val	Pro	Ala	Glu	Val	Ala	Lys	Val	Leu	Ala	Ser	Gly	Asn	Trp	
				1325					1330					1335		
acc	cca	gag	tct	cca	tcc	cca	gcc	tgc	cag	tgt	agc	cag	ccc	ggt	gcc	4265
Thr	Pro	Glu	Ser	Pro	Ser	Pro	Ala	Cys	Gln	Cys	Ser	Gln	Pro	Gly	Ala	
			1340					1345					1350			
cgg	cgc	ctg	ctg	ccc	gac	tgc	ccg	gct	gca	gct	ggt	ggt	ccc	cct	ccg	4313
Arg	Arg	Leu	Leu	Pro	Asp	Cys	Pro	Ala	Ala	Ala	Gly	Gly	Pro	Pro	Pro	
		1355					1360					1365				
ccc	cag	gca	gtg	acc	ggc	tct	ggg	gaa	gtg	gtt	cag	aac	ctg	aca	ggc	4361
Pro	Gln	Ala	Val	Thr	Gly	Ser	Gly	Glu	Val	Val	Gln	Asn	Leu	Thr	Gly	
		1370				1375					1380					
cgg	aac	ctg	tct	gac	ttc	ctg	gtc	aag	acc	tac	ccg	cgc	ctg	gtg	cgc	4409
Arg	Asn	Leu	Ser	Asp	Phe	Leu	Val	Lys	Thr	Tyr	Pro	Arg	Leu	Val	Arg	
1385					1390					1395					1400	
cag	ggc	ctg	aag	act	aag	aag	tgg	gtg	aat	gag	gtc	agg	tac	gga	ggc	4457
Gln	Gly	Leu	Lys	Thr	Lys	Lys	Trp	Val	Asn	Glu	Val	Arg	Tyr	Gly	Gly	
			1405						1410					1415		
ttc	tcg	ctg	ggg	ggc	cga	gac	cca	ggc	ctg	ccc	tcg	ggc	caa	gag	ttg	4505
Phe	Ser	Leu	Gly	Gly	Arg	Asp	Pro	Gly	Leu	Pro	Ser	Gly	Gln	Glu	Leu	
			1420					1425					1430			

[illegible]

ggc	cgc	tca	gtg	gag	gag	ttg	tgg	gcg	ctg	ctg	agt	ccc	ctg	cct	ggc	4553
Gly	Arg	Ser	Val	Glu	Glu	Leu	Trp	Ala	Leu	Leu	Ser	Pro	Leu	Pro	Gly	
1435			1440					1445								
ggg	gcc	ctc	gac	cgt	gtc	ctg	aaa	aac	ctc	aca	gcc	tgg	gct	cac	agc	4601
Gly	Ala	Leu	Asp	Arg	Val	Leu	Lys	Asn	Leu	Thr	Ala	Trp	Ala	His	Ser	
1450			1455					1460								
ctg	gat	gct	cag	gac	agt	ctc	aag	atc	tgg	ttc	aac	aac	aaa	ggc	tgg	4649
Leu	Asp	Ala	Gln	Asp	Ser	Leu	Lys	Ile	Trp	Phe	Asn	Asn	Lys	Gly	Trp	
1465			1470					1475					1480			
cac	tcc	atg	gtg	gcc	ttt	gtc	aac	cga	gcc	agc	aac	gca	atc	ctc	cgt	4697
His	Ser	Met	Val	Ala	Phe	Val	Asn	Arg	Ala	Ser	Asn	Ala	Ile	Leu	Arg	
1485			1490					1495								
gct	cac	ctg	ccc	cca	ggc	ccg	gcc	cgc	cac	gcc	cac	agc	atc	acc	aca	4745
Ala	His	Leu	Pro	Pro	Gly	Pro	Ala	Arg	His	Ala	His	Ser	Ile	Thr	Thr	
1500			1505					1510								
ctc	aac	cac	ccc	ttg	aac	ctc	acc	aag	gag	cag	ctg	tct	gag	gct	gca	4793
Leu	Asn	His	Pro	Leu	Asn	Leu	Thr	Lys	Glu	Gln	Leu	Ser	Glu	Ala	Ala	
1515			1520					1525								
ctg	atg	gcc	tcc	tcg	gtg	gac	gtc	ctc	gtc	tcc	atc	tgt	gtg	gtc	ttt	4841
Leu	Met	Ala	Ser	Ser	Val	Asp	Val	Leu	Val	Ser	Ile	Cys	Val	Val	Phe	
1530			1535					1540								
gcc	atg	tcc	ttt	gtc	ccg	gcc	agc	ttc	act	ctt	gtc	ctc	att	gag	gag	4889
Ala	Met	Ser	Phe	Val	Pro	Ala	Ser	Phe	Thr	Leu	Val	Leu	Ile	Glu	Glu	
1545			1550					1555					1560			
cga	gtc	acc	cga	gcc	aag	cac	ctg	cag	ctc	atg	ggg	ggc	ctg	tcc	ccc	4937
Arg	Val	Thr	Arg	Ala	Lys	His	Leu	Gln	Leu	Met	Gly	Gly	Leu	Ser	Pro	
1565			1570					1575								
acc	ctc	tac	tgg	ctt	ggc	aac	ttt	ctc	tgg	gac	atg	tgt	aac	tac	ttg	4985
Thr	Leu	Tyr	Trp	Leu	Gly	Asn	Phe	Leu	Trp	Asp	Met	Cys	Asn	Tyr	Leu	
1580			1585					1590								
gtg	cca	gca	tgc	atc	gtg	gtg	ctc	atc	ttt	ctg	gcc	ttc	cag	cag	agg	5033
Val	Pro	Ala	Cys	Ile	Val	Val	Leu	Ile	Phe	Leu	Ala	Phe	Gln	Gln	Arg	
1595			1600					1605								
gca	tat	gtg	gcc	cct	gcc	aac	ctg	cct	gct	ctc	ctg	ctg	ttg	cta	cta	5081
Ala	Tyr	Val	Ala	Pro	Ala	Asn	Leu	Pro	Ala	Leu	Leu	Leu	Leu	Leu	Leu	
1610			1615					1620								
ctg	tat	ggc	tgg	tcg	atc	aca	ccg	ctc	atg	tac	cca	gcc	tcc	ttc	ttc	5129
Leu	Tyr	Gly	Trp	Ser	Ile	Thr	Pro	Leu	Met	Tyr	Pro	Ala	Ser	Phe	Phe	
1625			1630					1635					1640			

# FIG. 2I

ttc tcc gtg ccc agc aca gcc tat gtg gtg ctc acc tgc ata aac ctc	5177
Phe Ser Val Pro Ser Thr Ala Tyr Val Val Leu Thr Cys Ile Asn Leu	
1645 1650 1655	
ttt att ggc atc aat gga agc atg gcc acc ttt gtg ctt gag ctc ttc	5225
Phe Ile Gly Ile Asn Gly Ser Met Ala Thr Phe Val Leu Glu Leu Phe	
1660 1665 1670	
tct gat cag aag ctg cag gag gtg agc cgg atc ttg aaa cag gtc ttc	5273
Ser Asp Gln Lys Leu Gln Glu Val Ser Arg Ile Leu Lys Gln Val Phe	
1675 1680 1685	
ctt atc ttc ccc cac ttc tgc ttg ggc cgg ggg ctc att gac atg gtg	5321
Leu Ile Phe Pro His Phe Cys Leu Gly Arg Gly Leu Ile Asp Met Val	
1690 1695 1700	
cgg aac cag gcc atg gct gat gcc ttt gag cgc ttg gga gac agg cag	5369
Arg Asn Gln Ala Met Ala Asp Ala Phe Glu Arg Leu Gly Asp Arg Gln	
1705 1710 1715 1720	
ttc cag tca ccc ctg cgc tgg gag gtg gtc ggc aag aac ctc ttg gcc	5417
Phe Gln Ser Pro Leu Arg Trp Glu Val Val Gly Lys Asn Leu Leu Ala	
1725 1730 1735	
atg gtg ata cag ggg ccc ctc ttc ctt ctc ttc aca cta ctg ctg cag	5465
Met Val Ile Gln Gly Pro Leu Phe Leu Leu Phe Thr Leu Leu Leu Gln	
1740 1745 1750	
cac cga agc caa ctc ctg cca cag ccc agg gtg agg tct ctg cca ctc	5513
His Arg Ser Gln Leu Leu Pro Gln Pro Arg Val Arg Ser Leu Pro Leu	
1755 1760 1765	
ctg gga gag gag gac gag gat gta gcc cgt gaa cgg gag cgg gtg gtc	5561
Leu Gly Glu Glu Asp Glu Asp Val Ala Arg Glu Arg Glu Arg Val Val	
1770 1775 1780	
caa gga gcc acc cag ggg gat gtg ttg gtg ctg agg aac ttg acc aag	5609
Gln Gly Ala Thr Gln Gly Asp Val Leu Val Leu Arg Asn Leu Thr Lys	
1785 1790 1795 1800	
gta tac cgt ggg cag agg atg cca gct gtt gac cgc ttg tgc ctg ggg	5657
Val Tyr Arg Gly Gln Arg Met Pro Ala Val Asp Arg Leu Cys Leu Gly	
1805 1810 1815	
att ccc cct ggt gag tgt ttt ggg ctg ctg ggt gtg aat gga gca ggg	5705
Ile Pro Pro Gly Glu Cys Phe Gly Leu Leu Gly Val Asn Gly Ala Gly	
1820 1825 1830	
aag acg tcc acg ttt cgc atg gtg acg ggg gac aca ttg gcc agc agg	5753
Lys Thr Ser Thr Phe Arg Met Val Thr Gly Asp Thr Leu Ala Ser Arg	
1835 1840 1845	

[illegible]

ggc gag gct gtg ctc gca ggc cac agc gtg gcc cgg gaa ccc agt gct	5801
Gly Glu Ala Val Leu Ala Gly His Ser Val Ala Arg Glu Pro Ser Ala	
1850 1855 1860	
gcg cac ctc agc atg gga tac tgc cct caa tcc gat gcc atc ttt gag	5849
Ala His Leu Ser Met Gly Tyr Cys Pro Gln Ser Asp Ala Ile Phe Glu	
1865 1870 1875 1880	
ctg ctg acg ggc cgc gag cac ctg gag ctg ctt gcg cgc ctg cgc ggt	5897
Leu Leu Thr Gly Arg Glu His Leu Glu Leu Leu Ala Arg Leu Arg Gly	
1885 1890 1895	
gtc ccg gag gcc cag gtt gcc cag acc gct ggc tca ggc ctg gcg cgt	5945
Val Pro Glu Ala Gln Val Ala Gln Thr Ala Gly Ser Gly Leu Ala Arg	
1900 1905 1910	
ctg gga ctc tca tgg tac gca gac cgg cct gca ggc acc tac agc gga	5993
Leu Gly Leu Ser Trp Tyr Ala Asp Arg Pro Ala Gly Thr Tyr Ser Gly	
1915 1920 1925	
ggg aac aaa cgc aag ctg gcg acg gcc ctg gcg ctg gtt ggg gac cca	6041
Gly Asn Lys Arg Lys Leu Ala Thr Ala Leu Ala Leu Val Gly Asp Pro	
1930 1935 1940	
gcc gtg gtg ttt ctg gac gag ccg acc aca ggc atg gac ccc agc gcg	6089
Ala Val Val Phe Leu Asp Glu Pro Thr Thr Gly Met Asp Pro Ser Ala	
1945 1950 1955 1960	
cgg cgc ttc ctt tgg aac agc ctt ttg gcc gtg gtg cgg gag ggc cgt	6137
Arg Arg Phe Leu Trp Asn Ser Leu Leu Ala Val Val Arg Glu Gly Arg	
1965 1970 1975	
tca gtg atg ctc acc tcc cat agc atg gag gag tgt gaa gcg ctc tgc	6185
Ser Val Met Leu Thr Ser His Ser Met Glu Glu Cys Glu Ala Leu Cys	
1980 1985 1990	
tcg cgc cta gcc atc atg gtg aat ggg cgg ttc cgc tgc ctg ggc agc	6233
Ser Arg Leu Ala Ile Met Val Asn Gly Arg Phe Arg Cys Leu Gly Ser	
1995 2000 2005	
ccg caa cat ctc aag ggc aga ttc gcg gcg ggt cac aca ctg acc ctg	6281
Pro Gln His Leu Lys Gly Arg Phe Ala Ala Gly His Thr Leu Thr Leu	
2010 2015 2020	
cgg gtg ccc gcc gca agg tcc cag ccg gca gcg gcc ttc gtg gcg gcc	6329
Arg Val Pro Ala Ala Arg Ser Gln Pro Ala Ala Ala Phe Val Ala Ala	
2025 2030 2035 2040	
gag ttc cct ggg tcg gag ctg cgc gag gca cat gga ggc cgc ctg cgc	6377
Glu Phe Pro Gly Ser Glu Leu Arg Glu Ala His Gly Gly Arg Leu Arg	
2045 2050 2055	

Parameter	Value	Unit
Temperature	25	°C
Pressure	1.0	atm
Flow rate	1.0	L/min
Concentration	0.1	mol/L
pH	7.0	
Time	10	min
Wavelength	254	nm
Scan rate	10	nm/min
Resolution	0.1	nm
Integration time	1.0	s
Detector gain	1.0	
Beam size	1.0	mm
Slit width	1.0	mm
Slit height	1.0	mm
Slit depth	1.0	mm
Slit angle	1.0	°
Slit position	1.0	mm
Slit shape	1.0	mm
Slit material	1.0	mm
Slit color	1.0	mm
Slit texture	1.0	mm
Slit weight	1.0	mm
Slit length	1.0	mm
Slit width	1.0	mm
Slit height	1.0	mm
Slit depth	1.0	mm
Slit angle	1.0	°
Slit position	1.0	mm
Slit shape	1.0	mm
Slit material	1.0	mm
Slit color	1.0	mm
Slit texture	1.0	mm
Slit weight	1.0	mm
Slit length	1.0	mm
Slit width	1.0	mm
Slit height	1.0	mm
Slit depth	1.0	mm
Slit angle	1.0	°
Slit position	1.0	mm
Slit shape	1.0	mm
Slit material	1.0	mm
Slit color	1.0	mm
Slit texture	1.0	mm
Slit weight	1.0	mm
Slit length	1.0	mm
Slit width	1.0	mm
Slit height	1.0	mm
Slit depth	1.0	mm
Slit angle	1.0	°
Slit position	1.0	mm
Slit shape	1.0	mm
Slit material	1.0	mm
Slit color	1.0	mm
Slit texture	1.0	mm
Slit weight	1.0	mm
Slit length	1.0	mm
Slit width	1.0	mm
Slit height	1.0	mm
Slit depth	1.0	mm
Slit angle	1.0	°
Slit position	1.0	mm
Slit shape	1.0	mm
Slit material	1.0	mm
Slit color	1.0	mm
Slit texture	1.0	mm
Slit weight	1.0	mm
Slit length	1.0	mm
Slit width	1.0	mm
Slit height	1.0	mm
Slit depth	1.0	mm
Slit angle	1.0	°
Slit position	1.0	mm
Slit shape	1.0	mm
Slit material	1.0	mm
Slit color	1.0	mm
Slit texture	1.0	mm
Slit weight	1.0	mm
Slit length	1.0	mm
Slit width	1.0	mm
Slit height	1.0	mm
Slit depth	1.0	mm
Slit angle	1.0	°
Slit position	1.0	mm
Slit shape	1.0	mm
Slit material	1.0	mm
Slit color	1.0	mm
Slit texture	1.0	mm
Slit weight	1.0	mm
Slit length	1.0	mm
Slit width	1.0	mm
Slit height	1.0	mm
Slit depth	1.0	mm
Slit angle	1.0	°
Slit position	1.0	mm
Slit shape	1.0	mm
Slit material	1.0	mm
Slit color	1.0	mm
Slit texture	1.0	mm
Slit weight	1.0	mm
Slit length	1.0	mm
Slit width	1.0	mm
Slit height	1.0	mm
Slit depth	1.0	mm
Slit angle	1.0	°
Slit position	1.0	mm
Slit shape	1.0	mm
Slit material	1.0	mm
Slit color	1.0	mm
Slit texture	1.0	mm
Slit weight	1.0	mm
Slit length	1.0	mm
Slit width	1.0	mm
Slit height	1.0	mm
Slit depth	1.0	mm
Slit angle	1.0	°
Slit position	1.0	mm
Slit shape	1.0	mm
Slit material	1.0	mm
Slit color	1.0	mm
Slit texture	1.0	mm
Slit weight	1.0	mm
Slit length	1.0	mm
Slit width	1.0	mm
Slit height	1.0	mm
Slit depth	1.0	mm
Slit angle	1.0	°
Slit position	1.0	mm
Slit shape</		

ttc	cag	ctg	ccg	ccg	gga	ggg	cgc	tgc	gcc	ctg	gcg	cgc	gtc	ttt	gga	6425
Phe	Gln	Leu	Pro	Pro	Gly	Gly	Arg	Cys	Ala	Leu	Ala	Arg	Val	Phe	Gly	
			2060			2065						2070				
gag	ctg	gcg	gtg	cac	ggc	gca	gag	cac	ggc	gtg	gag	gac	ttt	tcc	gtg	6473
Glu	Leu	Ala	Val	His	Gly	Ala	Glu	His	Gly	Val	Glu	Asp	Phe	Ser	Val	
			2075			2080						2085				
agc	cag	acg	atg	ctg	gag	gag	gta	ttc	ttg	tac	ttc	tcc	aag	gac	cag	6521
Ser	Gln	Thr	Met	Leu	Glu	Glu	Val	Phe	Leu	Tyr	Phe	Ser	Lys	Asp	Gln	
			2090			2095						2100				
ggg	aag	gac	gag	gac	acc	gaa	gag	cag	aag	gag	gca	gga	gtg	gga	gtg	6569
Gly	Lys	Asp	Glu	Asp	Thr	Glu	Glu	Gln	Lys	Glu	Ala	Gly	Val	Gly	Val	
2105			2110						2115			2120				
gac	ccc	gcg	cca	ggc	ctg	cag	cac	ccc	aaa	cgc	gtc	agc	cag	ttc	ctc	6617
Asp	Pro	Ala	Pro	Gly	Leu	Gln	His	Pro	Lys	Arg	Val	Ser	Gln	Phe	Leu	
			2125						2130			2135				
gat	gac	cct	agc	act	gcc	gag	act	gtg	ctc	tga	gcctccctcc	cctgcggggc				6670
Asp	Asp	Pro	Ser	Thr	Ala	Glu	Thr	Val	Leu							
			2140			2145										
cgcggggagg		ccctgggaat		ggcaagggca		aggtagagtg		cctaggagcc		ctggactcag		6730				
gctggcagag		gggctggtgc		cctggagaaa		ataaagagaa		ggctggagag		aagccgtgct		6790				
ggtgaaaaaa		aaaa										6804				

# FIG. 3A

atg gtc tgc ctg gga act ggc cag agc gct gga ccc cta gtg agt gtt	48
Met Val Cys Leu Gly Thr Gly Gln Ser Ala Gly Pro Leu Val Ser Val	
1 5 10 15	
caa aat cat tgt ccc cct tgt ggt ctt tct ccc cag gaa tcc ctg ggg	96
Gln Asn His Cys Pro Pro Cys Gly Leu Ser Pro Gln Glu Ser Leu Gly	
20 25 30	
ttg gca ctg ggc caa gcc cag gag ccc ttg cac agc ttg ttg gag gcc	144
Leu Ala Leu Gly Gln Ala Gln Glu Pro Leu His Ser Leu Leu Glu Ala	
35 40 45	
gct ggg gac ctg gcc cag gag ctc ctg gcg ctg cgc agc ctg gtg gag	192
Ala Gly Asp Leu Ala Gln Glu Leu Leu Ala Leu Arg Ser Leu Val Glu	
50 55 60	
ctt cgg gca ctg ctg cag aga ccc cga ggg acc agc ggc ccc ctg gag	240
Leu Arg Ala Leu Leu Gln Arg Pro Arg Gly Thr Ser Gly Pro Leu Glu	
65 70 75 80	
ttg ctg tca gag gcc ctc tgc agt gtc agg gga cct agc agc aca gtg	288
Leu Leu Ser Glu Ala Leu Cys Ser Val Arg Gly Pro Ser Ser Thr Val	
85 90 95	
ggc ccc tcc ctc aac tgg tac gag gct agt gac ctg atg gag ctg gtg	336
Gly Pro Ser Leu Asn Trp Tyr Glu Ala Ser Asp Leu Met Glu Leu Val	
100 105 110	
ggg cag gag cca gaa tcc gcc ctg cca gac agc agc ctg agc ccc gcc	384
Gly Gln Glu Pro Glu Ser Ala Leu Pro Asp Ser Ser Leu Ser Pro Ala	
115 120 125	
tgc tcg gag ctg att gga gcc ctg gac agc cac ccg ctg tcc cgc ctg	432
Cys Ser Glu Leu Ile Gly Ala Leu Asp Ser His Pro Leu Ser Arg Leu	
130 135 140	
ctc tgg aga cgc ctg aag cct ctg atc ctc ggg aag cta ctc ttt gca	480
Leu Trp Arg Arg Leu Lys Pro Leu Ile Leu Gly Lys Leu Leu Phe Ala	
145 150 155 160	
cca gat aca cct ttt acc cgg aag ctc atg gcc cag gtg aac cgg acc	528
Pro Asp Thr Pro Phe Thr Arg Lys Leu Met Ala Gln Val Asn Arg Thr	
165 170 175	
ttc gag gag ctc acc ctg ctg agg gat gtc cgg gag gtg tgg gag atg	576
Phe Glu Glu Leu Thr Leu Leu Arg Asp Val Arg Glu Val Trp Glu Met	
180 185 190	
ctg gga ccc cgg atc ttc acc ttc atg aac gac agt tcc aat gtg gcc	624
Leu Gly Pro Arg Ile Phe Thr Phe Met Asn Asp Ser Ser Asn Val Ala	
195 200 205	

Sequence 1

FIG. 3B

atg ctg cag cgg ctc ctg cag atg cag gat gaa gga aga agg cag ccc	672
Met Leu Gln Arg Leu Leu Gln Met Gln Asp Glu Gly Arg Arg Gln Pro	
210 215 220	
aga cct gga ggc cgg gac cac atg gag gcc ctg cga tcc ttt ctg gac	720
Arg Pro Gly Gly Arg Asp His Met Glu Ala Leu Arg Ser Phe Leu Asp	
225 230 235 240	
cct ggg agc ggt ggc tac agc tgg cag gac gca cac gct gat gtg ggg	768
Pro Gly Ser Gly Gly Tyr Ser Trp Gln Asp Ala His Ala Asp Val Gly	
245 250 255	
cac ctg gtg ggc acg ctg ggc cga gtg acg gag tgc ctg tcc ttg gac	816
His Leu Val Gly Thr Leu Gly Arg Val Thr Glu Cys Leu Ser Leu Asp	
260 265 270	
aag ctg gag gcg gca ccc tca gag gca gcc ctg gtg tcg cgg gcc ctg	864
Lys Leu Glu Ala Ala Pro Ser Glu Ala Ala Leu Val Ser Arg Ala Leu	
275 280 285	
caa ctg ctc gcg gaa cat cga ttc tgg gcc ggc gtc gtc ttc ttg gga	912
Gln Leu Leu Ala Glu His Arg Phe Trp Ala Gly Val Val Phe Leu Gly	
290 295 300	
cct gag gac tct tca gac ccc aca gag cac cca acc cca gac ctg ggc	960
Pro Glu Asp Ser Ser Asp Pro Thr Glu His Pro Thr Pro Asp Leu Gly	
305 310 315 320	
ccc ggc cac gtg cgc atc aaa atc cgc atg gac att gac gtg gtc acg	1008
Pro Gly His Val Arg Ile Lys Ile Arg Met Asp Ile Asp Val Val Thr	
325 330 335	
agg acc aat aag atc agg gac agg ttt tgg gac cct ggc cca gcc gcg	1056
Arg Thr Asn Lys Ile Arg Asp Arg Phe Trp Asp Pro Gly Pro Ala Ala	
340 345 350	
gac ccc ctg acc gac ctg cgc tac gtg tgg ggc ggc ttc gtg tac ctg	1104
Asp Pro Leu Thr Asp Leu Arg Tyr Val Trp Gly Gly Phe Val Tyr Leu	
355 360 365	
caa gac ctg gtg gag cgt gca gcc gtc cgc gtg ctc agc ggc gcc aac	1152
Gln Asp Leu Val Glu Arg Ala Ala Val Arg Val Leu Ser Gly Ala Asn	
370 375 380	
ccc cgg gcc ggc ctc tac ctg cag cag atg ccc tat ccg tgc tat gtg	1200
Pro Arg Ala Gly Leu Tyr Leu Gln Gln Met Pro Tyr Pro Cys Tyr Val	
385 390 395 400	
gac gac gtg ttc ctg cgt gtg ctg agc cgg tcg ctg ccg ctc ttc ctg	1248
Asp Asp Val Phe Leu Arg Val Leu Ser Arg Ser Leu Pro Leu Phe Leu	
405 410 415	



FIG. 3C

acg ctg gcc tgg atc tac tcc gtg aca ctg aca gtg aag gcc gtg gtg	1296
Thr Leu Ala Trp Ile Tyr Ser Val Thr Leu Thr Val Lys Ala Val Val	
420 425 430	
cgg gag aag gag acg cgg ctg cgg gac acc atg cgc gcc atg ggg ctc	1344
Arg Glu Lys Glu Thr Arg Leu Arg Asp Thr Met Arg Ala Met Gly Leu	
435 440 445	
agc cgc gcg gtg ctc tgg cta ggc tgg ttc ctc agc tgc ctc ggg ccc	1392
Ser Arg Ala Val Leu Trp Leu Gly Trp Phe Leu Ser Cys Leu Gly Pro	
450 455 460	
ttc ctg ctc agc gcc gcg ctg ctg gtt ctg gtg ctc aag ctg ggg gac	1440
Phe Leu Leu Ser Ala Ala Leu Leu Val Leu Val Leu Lys Leu Gly Asp	
465 470 475 480	
atc ctc ccc tac agc cac ccg ggc gtg gtc ttc ctg ttc ttg gca gcc	1488
Ile Leu Pro Tyr Ser His Pro Gly Val Val Phe Leu Phe Leu Ala Ala	
485 490 495	
ttc gcg gtg gcc acg gtg acc cag agc ttc ctg ctc agc gcc ttc ttc	1536
Phe Ala Val Ala Thr Val Thr Gln Ser Phe Leu Leu Ser Ala Phe Phe	
500 505 510	
tcc cgc gcc aac ctg gct gcg gcc tgc ggc ggc ctg gcc tac ttc tcc	1584
Ser Arg Ala Asn Leu Ala Ala Ala Cys Gly Gly Leu Ala Tyr Phe Ser	
515 520 525	
ctc tac ctg ccc tac gtg ctg tgt gtg gct tgg cgg gac cgg ctg ccc	1632
Leu Tyr Leu Pro Tyr Val Leu Cys Val Ala Trp Arg Asp Arg Leu Pro	
530 535 540	
gcg ggt ggc cgc gtg gcc gcg agc ctg ctg tcg ccc gtg gcc ttc ggc	1680
Ala Gly Gly Arg Val Ala Ala Ser Leu Leu Ser Pro Val Ala Phe Gly	
545 550 555 560	
ttc ggc tgc gag agc ctg gct ctg ctg gag gag cag ggc gag ggc gcg	1728
Phe Gly Cys Glu Ser Leu Ala Leu Leu Glu Glu Gln Gly Glu Gly Ala	
565 570 575	
cag tgg cac aac gtg ggc acc cgg cct acg gca gac gtc ttc agc ctg	1776
Gln Trp His Asn Val Gly Thr Arg Pro Thr Ala Asp Val Phe Ser Leu	
580 585 590	
gcc cag gtc tct ggc ctt ctg ctg ctg gac gcg gcg ctc tac ggc ctc	1824
Ala Gln Val Ser Gly Leu Leu Leu Leu Asp Ala Ala Leu Tyr Gly Leu	
595 600 605	
gcc acc tgg tac ctg gaa gct gtg tgc cca ggc cag tac ggg atc cct	1872
Ala Thr Trp Tyr Leu Glu Ala Val Cys Pro Gly Gln Tyr Gly Ile Pro	
610 615 620	

# FIG. 3D

gaa cca tgg aat ttt cct ttt cgg agg agc tac tgg tgc gga cct cgg	1920
Glu Pro Trp Asn Phe Pro Phe Arg Arg Ser Tyr Trp Cys Gly Pro Arg	
625 630 635 640	
ccc ccc aag agt cca gcc cct tgc ccc acc ccg ctg gac cca aag gtg	1968
Pro Pro Lys Ser Pro Ala Pro Cys Pro Thr Pro Leu Asp Pro Lys Val	
645 650 655	
ctg gta gaa gag gca ccg ccc ggc ctg agt cct ggc gta tcc gtt cgc	2016
Leu Val Glu Glu Ala Pro Pro Gly Leu Ser Pro Gly Val Ser Val Arg	
660 665 670	
agc ctg gag aag cgc ttt cct gga agc ccg cag cca gcc ctg cgg ggg	2064
Ser Leu Glu Lys Arg Phe Pro Gly Ser Pro Gln Pro Ala Leu Arg Gly	
675 680 685	
ctc agc ctg gac ttc tac cag ggc cac atc acc gcc ttc ctg ggc cac	2112
Leu Ser Leu Asp Phe Tyr Gln Gly His Ile Thr Ala Phe Leu Gly His	
690 695 700	
aac ggg gcc ggc aag acc acc acc ctg tcc atc ttg agt ggc ctc ttc	2160
Asn Gly Ala Gly Lys Thr Thr Leu Ser Ile Leu Ser Gly Leu Phe	
705 710 715 720	
cca ccc agt ggt ggc tct gcc ttc atc ctg ggc cac gac gtc cgc tcc	2208
Pro Pro Ser Gly Gly Ser Ala Phe Ile Leu Gly His Asp Val Arg Ser	
725 730 735	
agc atg gcc gcc atc cgg ccc cac ctg ggc gtc tgt cct cag tac aac	2256
Ser Met Ala Ala Ile Arg Pro His Leu Gly Val Cys Pro Gln Tyr Asn	
740 745 750	
gtg ctg ttt gac atg ctg acc gtg gac gag cac gtc tgg ttc tat ggg	2304
Val Leu Phe Asp Met Leu Thr Val Asp Glu His Val Trp Phe Tyr Gly	
755 760 765	
cgg ctg aag ggt ctg agt gcc gct gta gtg ggc ccc gag cag gac cgt	2352
Arg Leu Lys Gly Leu Ser Ala Ala Val Val Gly Pro Glu Gln Asp Arg	
770 775 780	
ctg ctg cag gat gtg ggg ctg gtc tcc aag cag agt gtg cag act cgc	2400
Leu Leu Gln Asp Val Gly Leu Val Ser Lys Gln Ser Val Gln Thr Arg	
785 790 795 800	
cac ctc tct ggt ggg atg caa cgg aag ctg tcc gtg gcc att gcc ttt	2448
His Leu Ser Gly Gly Met Gln Arg Lys Leu Ser Val Ala Ile Ala Phe	
805 810 815	
gtg ggc ggc tcc caa gtt gtt atc ctg gac gag cct acg gct ggc gtg	2496
Val Gly Gly Ser Gln Val Val Ile Leu Asp Glu Pro Thr Ala Gly Val	
820 825 830	

FIG. 3E

gat cct gct tcc cgc cgc ggt att tgg gag ctg ctg ctc aaa tac cga	2544
Asp Pro Ala Ser Arg Arg Gly Ile Trp Glu Leu Leu Leu Lys Tyr Arg	
835 840 845	
gaa ggt cgc acg ctg atc ctc tcc acc cac cac ctg gat gag gca gag	2592
Glu Gly Arg Thr Leu Ile Leu Ser Thr His His Leu Asp Glu Ala Glu	
850 855 860	
ctg ctg gga gac cgt gtg gct gtg gtg gca ggt ggc cgc ttg tgc tgc	2640
Leu Leu Gly Asp Arg Val Ala Val Val Ala Gly Gly Arg Leu Cys Cys	
865 870 875 880	
tgt ggc tcc cca ctc ttc ctg cgc cgt cac ctg ggc tcc ggc tac tac	2688
Cys Gly Ser Pro Leu Phe Leu Arg Arg His Leu Gly Ser Gly Tyr Tyr	
885 890 895	
ctg acg ctg gtg aag gcc cgc ctg ccc ctg acc acc aat gag aag gct	2736
Leu Thr Leu Val Lys Ala Arg Leu Pro Leu Thr Thr Asn Glu Lys Ala	
900 905 910	
gac act gac atg gag ggc agt gtg gac acc agg cag gaa aag aag aat	2784
Asp Thr Asp Met Glu Gly Ser Val Asp Thr Arg Gln Glu Lys Lys Asn	
915 920 925	
ggc agc cag ggc agc aga gtc ggc act cct cag ctg ctg gcc ctg gta	2832
Gly Ser Gln Gly Ser Arg Val Gly Thr Pro Gln Leu Leu Ala Leu Val	
930 935 940	
cag cac tgg gtg ccc ggg gca cgg ctg gtg gag gag ctg cca cac gag	2880
Gln His Trp Val Pro Gly Ala Arg Leu Val Glu Glu Leu Pro His Glu	
945 950 955 960	
ctg gtg ctg gtg ctg ccc tac acg ggt gcc cat gac ggc agc ttc gcc	2928
Leu Val Leu Val Leu Pro Tyr Thr Gly Ala His Asp Gly Ser Phe Ala	
965 970 975	
aca ctc ttc cga gag cta gac acg cgg ctg gcg gag ctg agg ctc act	2976
Thr Leu Phe Arg Glu Leu Asp Thr Arg Leu Ala Glu Leu Arg Leu Thr	
980 985 990	
ggc tac ggg atc tcc gac acc agc ctc gag gag atc ttc ctg aag gtg	3024
Gly Tyr Gly Ile Ser Asp Thr Ser Leu Glu Glu Ile Phe Leu Lys Val	
995 1000 1005	
gtg gag gag tgt gct gcg gac aca gat atg gag gat ggc agc tgc ggg	3072
Val Glu Glu Cys Ala Ala Asp Thr Asp Met Glu Asp Gly Ser Cys Gly	
1010 1015 1020	
cag cac cta tgc aca ggc att gct ggc cta gac gta acc ctg cgg ctc	3120
Gln His Leu Cys Thr Gly Ile Ala Gly Leu Asp Val Thr Leu Arg Leu	
1025 1030 1035 1040	

FIG. 3F

aag atg ccg cca cag gag aca gcg ctg gag aac ggg gaa cca gct ggg	3168
Lys Met Pro Pro Gln Glu Thr Ala Leu Glu Asn Gly Glu Pro Ala Gly	
1045 1050 1055	
tca gcc cca gag act gac cag ggc tct ggg cca gac gcc gtg ggc cgg	3216
Ser Ala Pro Glu Thr Asp Gln Gly Ser Gly Pro Asp Ala Val Gly Arg	
1060 1065 1070	
gta cag ggc tgg gca ctg acc cgc cag cag ctc cag gcc ctg ctt ctc	3264
Val Gln Gly Trp Ala Leu Thr Arg Gln Gln Leu Gln Ala Leu Leu Leu	
1075 1080 1085	
aag cgc ttt ctg ctt gcc cgc cgc agc cgc cgc ggc ctg ttc gcc cag	3312
Lys Arg Phe Leu Leu Ala Arg Arg Ser Arg Arg Gly Leu Phe Ala Gln	
1090 1095 1100	
atc gtg ctg cct gcc ctc ttt gtg ggc ctg gcc ctc gtg ttc agc ctc	3360
Ile Val Leu Pro Ala Leu Phe Val Gly Leu Ala Leu Val Phe Ser Leu	
1105 1110 1115 1120	
atc gtg cct cct ttc ggg cac tac ccg gct ctg cgg ctc agt ccc acc	3408
Ile Val Pro Pro Phe Gly His Tyr Pro Ala Leu Arg Leu Ser Pro Thr	
1125 1130 1135	
atg tac ggt gct cag gtg tcc ttc ttc agt gag gac gcc cca ggg gac	3456
Met Tyr Gly Ala Gln Val Ser Phe Phe Ser Glu Asp Ala Pro Gly Asp	
1140 1145 1150	
cct gga cgt gcc cgg ctg ctc gag gcg ctg ctg cag gag gca gga ctg	3504
Pro Gly Arg Ala Arg Leu Leu Glu Ala Leu Leu Gln Glu Ala Gly Leu	
1155 1160 1165	
gag gag ccc cca gtg cag cat agc tcc cac agg ttc tcg gca cca gaa	3552
Glu Glu Pro Pro Val Gln His Ser Ser His Arg Phe Ser Ala Pro Glu	
1170 1175 1180	
gtt cct gct gaa gtg gcc aag gtc ttg gcc agt ggc aac tgg acc cca	3600
Val Pro Ala Glu Val Ala Lys Val Leu Ala Ser Gly Asn Trp Thr Pro	
1185 1190 1195 1200	
gag tct cca tcc cca gcc tgc cag tgt agc cag ccc ggt gcc cgg cgc	3648
Glu Ser Pro Ser Pro Ala Cys Gln Cys Ser Gln Pro Gly Ala Arg Arg	
1205 1210 1215	
ctg ctg ccc gac tgc ccg gct gca gct ggt ggt ccc cct ccg ccc cag	3696
Leu Leu Pro Asp Cys Pro Ala Ala Ala Gly Gly Pro Pro Pro Gln	
1220 1225 1230	
gca gtg acc ggc tct ggg gaa gtg gtt cag aac ctg aca ggc cgg aac	3744
Ala Val Thr Gly Ser Gly Glu Val Val Gln Asn Leu Thr Gly Arg Asn	
1235 1240 1245	

FIG. 3G

ctg tct gac ttc ctg gtc aag acc tac ccg cgc ctg gtg cgc cag ggc	3792
Leu Ser Asp Phe Leu Val Lys Thr Tyr Pro Arg Leu Val Arg Gln Gly	
1250 1255 1260	
ctg aag act aag aag tgg gtg aat gag gtc agg tac gga ggc ttc tcg	3840
Leu Lys Thr Lys Lys Trp Val Asn Glu Val Arg Tyr Gly Gly Phe Ser	
1265 1270 1275 1280	
ctg ggg ggc cga gac cca ggc ctg ccc tcg ggc caa gag ttg ggc cgc	3888
Leu Gly Gly Arg Asp Pro Gly Leu Pro Ser Gly Gln Glu Leu Gly Arg	
1285 1290 1295	
tca gtg gag gag ttg tgg gcg ctg ctg agt ccc ctg cct ggc ggg gcc	3936
Ser Val Glu Glu Leu Trp Ala Leu Leu Ser Pro Leu Pro Gly Gly Ala	
1300 1305 1310	
ctc gac cgt gtc ctg aaa aac ctc aca gcc tgg gct cac agc ctg gat	3984
Leu Asp Arg Val Leu Lys Asn Leu Thr Ala Trp Ala His Ser Leu Asp	
1315 1320 1325	
gct cag gac agt ctc aag atc tgg ttc aac aac aaa ggc tgg cac tcc	4032
Ala Gln Asp Ser Leu Lys Ile Trp Phe Asn Asn Lys Gly Trp His Ser	
1330 1335 1340	
atg gtg gcc ttt gtc aac cga gcc agc aac gca atc ctc cgt gct cac	4080
Met Val Ala Phe Val Asn Arg Ala Ser Asn Ala Ile Leu Arg Ala His	
1345 1350 1355 1360	
ctg ccc cca ggc ccg gcc cgc cac gcc cac agc atc acc aca ctc aac	4128
Leu Pro Pro Gly Pro Ala Arg His Ala His Ser Ile Thr Thr Leu Asn	
1365 1370 1375	
cac ccc ttg aac ctc acc aag gag cag ctg tct gag gct gca ctg atg	4176
His Pro Leu Asn Leu Thr Lys Glu Gln Leu Ser Glu Ala Ala Leu Met	
1380 1385 1390	
gcc tcc tcg gtg gac gtc ctc gtc tcc atc tgt gtg gtc ttt gcc atg	4224
Ala Ser Ser Val Asp Val Leu Val Ser Ile Cys Val Val Phe Ala Met	
1395 1400 1405	
tcc ttt gtc ccg gcc agc ttc act ctt gtc ctc att gag gag cga gtc	4272
Ser Phe Val Pro Ala Ser Phe Thr Leu Val Leu Ile Glu Glu Arg Val	
1410 1415 1420	
acc cga gcc aag cac ctg cag ctc atg ggg ggc ctg tcc ccc acc ctc	4320
Thr Arg Ala Lys His Leu Gln Leu Met Gly Gly Leu Ser Pro Thr Leu	
1425 1430 1435 1440	
tac tgg ctt ggc aac ttt ctc tgg gac atg tgt aac tac ttg gtg cca	4368
Tyr Trp Leu Gly Asn Phe Leu Trp Asp Met Cys Asn Tyr Leu Val Pro	
1445 1450 1455	

# FIG. 3H

gca tgc atc gtg gtg ctc atc ttt ctg gcc ttc cag cag agg gca tat	4416
Ala Cys Ile Val Val Leu Ile Phe Leu Ala Phe Gln Gln Arg Ala Tyr	
1460 1465 1470	
gtg gcc cct gcc aac ctg cct gct ctc ctg ctg ttg cta cta ctg tat	4464
Val Ala Pro Ala Asn Leu Pro Ala Leu Leu Leu Leu Leu Leu Tyr	
1475 1480 1485	
ggg aga cag gca gtt cca gtc acc cct gcg ctg gga ggt ggt cgg caa	4512
Gly Arg Gln Ala Val Pro Val Thr Pro Ala Leu Gly Gly Gly Arg Gln	
1490 1495 1500	
gaa cct ctt ggc cat ggt gat aca ggg gcc cct ctt cct tct ctt cac	4560
Glu Pro Leu Gly His Gly Asp Thr Gly Ala Pro Leu Pro Ser Leu His	
1505 1510 1515 1520	
act act gct gca gca ccg aag cca act cct gcc aca gcc cag ggt gag	4608
Thr Thr Ala Ala Ala Pro Lys Pro Thr Pro Ala Thr Ala Gln Gly Glu	
1525 1530 1535	
gtc tct gcc act cct ggg aga gga gga cga gga tgt agc ccg tga	4653
Val Ser Ala Thr Pro Gly Arg Gly Gly Arg Gly Cys Ser Pro	
1540 1545 1550	

gca tgc atc gtg gtg ctc atc ttt ctg gcc ttc cag cag agg gca tat  
Ala Cys Ile Val Val Leu Ile Phe Leu Ala Phe Gln Gln Arg Ala Tyr  
1460 1465 1470  
gtg gcc cct gcc aac ctg cct gct ctc ctg ctg ttg cta cta ctg tat  
Val Ala Pro Ala Asn Leu Pro Ala Leu Leu Leu Leu Leu Leu Tyr  
1475 1480 1485  
ggg aga cag gca gtt cca gtc acc cct gcg ctg gga ggt ggt cgg caa  
Gly Arg Gln Ala Val Pro Val Thr Pro Ala Leu Gly Gly Gly Arg Gln  
1490 1495 1500  
gaa cct ctt ggc cat ggt gat aca ggg gcc cct ctt cct tct ctt cac  
Glu Pro Leu Gly His Gly Asp Thr Gly Ala Pro Leu Pro Ser Leu His  
1505 1510 1515 1520  
act act gct gca gca ccg aag cca act cct gcc aca gcc cag ggt gag  
Thr Thr Ala Ala Ala Pro Lys Pro Thr Pro Ala Thr Ala Gln Gly Glu  
1525 1530 1535  
gtc tct gcc act cct ggg aga gga gga cga gga tgt agc ccg tga  
Val Ser Ala Thr Pro Gly Arg Gly Gly Arg Gly Cys Ser Pro  
1540 1545 1550

FIG. 4A

	1		50
muABC1	.....	.....	.....
muABCR	MGFLRQIQLL	LWKNWTLRKR	QKIRFVVELV WPLSLFLVLI WLRNANPLYS
muABCL	MAFCTQLMLL	LWKNYTYRRR	QPIQLLVELL WPLFLFFILV AVRHSHPPLE
	51		100
muABC1	.....	MPSAGTLPWV	QGIICNANNP CRYPTOPTGEA PGVVGNFNKS
muABCR	QHECHFPNKA	MPSAGLLPWL	QGIFCNMNNP CFQNPTPGES PGTVSNYNNS
muABCL	HHECHFPNKP	LPSAGTVPWL	QGLVCNVNNS CFQHPTPGEK PGVLSNFKDS
	101		150
muABC1	IVSRLFSDAQ	RLLLYSQRDT	SIKDMHKVLR MLRQIKHPN. ....SN
muABCR	ILARVYRDFQ	ELFMDTPEVQ	HLGQVWAE LR TLSQFMDTLR THPERFAGRG
muABCL	LISRLADTR	TVLGGHSIQD	MLDALGKLIP VLRVGGGA. ....
	151		200
muABC1	LKLQDFLVDN	ETFSGFLQHN	LSLPRSTVDS LLQXNVGLQK VFLQGYQLHL
muABCR	LQIRDILKDE	EALTLFLMRN	IGLSDSVAHL LVNSQVRVEQ FAYGVPDLEL
muABCL	.....	.....	.....
	201		250
muABC1	ASLCNGSKLE	EIIQLG....	..DAEVSALC GLPRKKLDAA ERVLRYNMDI
muABCR	TDIACSEALL	QRFIIFSQRR	GAQTVRDALC PLSQVTLQWI EDTLYADVDF
muABCL	.....	.....	..RPQESDQP TSQGSVTKLL EKILQRASLD
	251		300
muABC1	LKPVVTKLNS	TSHLPTQHLA	EATTVLLDSL GGLAQELFST KSWSDMRQEV
muABCR	FKLFHVLPTL	LDSSSQGINL	RFWGGILSDL SPRMQKFIHR PSVQDLLWVS
muABCL	PVLGQAQDSM	RKFSDAIRDL	AQELLTLPSL MELRALLR.. R.....
	301		350
muABC1	MFLTNVNSSS	SSTQIYQAVS	RIVCGHPEGG GLKIKSLNWX EDNNYKALFG
muABCR	RPLLQNGGPE	TFTQLMSILS	DLLCGYPEGG GSRVFSFNWX EDNNYKAFLG
muABCL	.....PR	GSAGSLELVS	EALCSTKGPS SPGGLSLNWX EANQLNEFMG
	351		400
muABC1	GNNTEEDVDT	FYDNSTTPYC	NDLMKNLESS PLSRIIWKAL KPLLVGKILY
muABCR	IDSTRKDPAY	SYDKRTTSFC	NSLIQSLESN PLTKIAWRAA KPLLMGKILF
muABCL	....PEVAPA	LPDNLSPAC	SEFVGTLDDH PVSRLLRRL KPLILGKILF
	401		450
muABC1	TPDTPATRQV	MAEVNKTQFE	LAVFHDLEGM WEELSPQIWT FMENSQEMDL
muABCR	TPDSPAARRI	MKNANSTFEE	LDRVRKLVKA WEEVGPQIWX FFEKSTQMTV
muABCL	APDTNFTRKL	MAQVNQTFEE	LALLRDLHEL WGVLGPIFN FMNDSTNVAM
	451		500
muABC1	VRTLLDSRGN	DQFWEQKLDG	LDWTAQDIMA FLAKNPEDVQ SPNGSVYTWX
muABCR	IRDTLQHPTV	KDFINRQLGE	EGITTEAVLN FFSNGPQEKQ ADDMTSFDWR
muABCL	LQRLLDVGGT	GQRQQTGRAQ	.....K.. KL.EAIKDFL DPSRGGYSWR

Protein Data Bank

FIG. 4B

	501				550
muABC1	EAFNETNQAI	QTISRMECV	NLNKLEPIPT	EVRLINKSME	LLDERKFWAG
muABCR	DIFNITDRFL	RLANQYLECL	VLDKFESYDD	EVQLTQRALS	LLEENRFWAG
muABCL	EAHADMGRLA	GILGQMMECV	SLDKLEAVPS	EEALVSRAL	LLGERRLWAG
	551				600
muABC1	IVFTG.....	...ITPDSVE	LPHHVKYKIR	MDIDNVERTN	KIKDGYWDPG
muABCR	VVFPG.....	...MYPWASS	LPPHVKYKIR	MDIDVVEKTN	KIKDRYWDSG
muABCL	IVFLSPEHPL	DPSELSSPAL	SPGHLRFKIR	MDIDDVTRTN	KIRDKFWDPG
	601				650
muABC1	PRADPFEDMR	YVWGGFAYLQ	DVVEQAIIRV	LTGSEKKTGV	YVQQMPYPCY
muABCR	PRADPVEDFR	YIWGGFAYLQ	DMVEQGIKVS	QMQAEPPIGV	YLQQMPYPCF
muABCL	PSADPFMDLR	YVWGGFVYLQ	DLLEQAAVRV	LGGGNSRTGL	YLQQMPHPCY
	651				700
muABC1	VDDIFLRVMS	RSMPLFMTLA	WIYSVAVIK	SIVYEKEARL	KETMRIMGLD
muABCR	VDDSFMIILN	RCFPIFMVLA	WIYSVSMTVK	GIVLEKELRL	KETLKNQGV
muABCL	VDDVFLRVLS	RSLPLFLTLA	WIYSVALTVK	AVVREKETRL	RETMRAMGLS
	701				750
muABC1	NGILWFSWFV	SSLIPLLVSA	GLLVVILKLG	NLLPYSDPSV	VFVFLSVFAM
muABCR	NAVIWCTWFL	DSFSIMALSI	FLLTFLFIMHG	RILHYSDFPI	LFLFLLAFAF
muABCL	RAVLWLGWFL	SCLGPFLVSA	ALLVLVLKLG	NILPYSHPVV	IFLFLAAFAV
	751				800
muABC1	VTLQCFLIS	TLFSRANLAA	ACGGIIYFTL	YLPYVLCVAW	QDYVGFSIKI
muABCR	ATIMQSFLLS	TLFSKASLAA	ACSGVIYFTL	YLPVLCFAW	QDRMTADLKT
muABCL	ATVAQSFLLS	AFFSRANLAA	ACGGLAYFAL	YLPYVLCVAW	RERLHLGGLL
	801				850
muABC1	FASLLSPVAF	GFGCEYFALF	EEQGIGVQWD	NLFESPVEED	GFNLTTAVSM
muABCR	TVSLLSSVAF	GFGTEYLVR	EEQGLGLQWS	NIGKSPLEGD	EFSFLLSMKM
muABCL	AASLLSPVAF	GFGCESLALL	EEQGDGAQWH	NLGTG.PAED	VFSLAQVSAF
	851				900
muABC1	MLFDTFLYGV	MTWYIEAVFP	GQYGIPRPWY	FPCTKSYWFG	.....
muABCR	MLLDAALYGL	LAWYLDQVFP	GDYGTPLPWY	FLLQESYWL	GEGCSTREER
muABCL	LLLDAVIYGL	ALWYLEAVCP	GQYGIPEPWN	FPFRRSYWCG	.....
	901				950
muABC1	.....EEIDE	KSHPGSSQKG	VSEICMEEEP	THLRLGVSIQ	NLVKVYRDGM
muABCR	ALEKTEPLTE	EMEDPEHPEG	MNDSFFEREL	PGLVPGVCVK	NLVKVFEPSP
muABCL	.....PGP.P	KSSVLAPAPQ	DPKVLVEEPP	LGLVPGVSIR	GLKKHFRGCP
	951				1000
muABC1	KVAVDGLALN	FYEGQITSFL	GHNGAGKTTT	MSILTGLFPP	TSGTAYILGK
muABCR	RPAVDRLNIT	FYENQITAFL	GHNGAGKTTT	LSILTGLLPP	TSGTVLIGGK
muABCL	QPALQGLNLD	FYEGHITAFL	GHNGAGKTTT	LSILSGLFPP	SSGSASILGH



FIG. 4C

	1001				1050
muABC1	DIRSEMSSIR	QNLGVCPQHN	VLFDMLTVEE	HIWFYARLKG	LSEKHVKAEM
muABCR	DIETNLDVVR	QSLGMCPQHN	ILFHHLTVAE	HILFYAQLKG	RSWEEAQLEM
muABCL	DVQTNMAAIR	PHLGICPQYN	VLFDMLTVEE	HVWFYGRLLKG	VSAAAMGPER
	1051				1100
muABC1	EQMALDVGLP	PSKLKSKTSQ	LSGGMQRKLS	VALAFVGGSK	VVILDEPTAG
muABCR	EAMLEDTGLH	HK.RNEEAQD	LSGGMQRKLS	VAIAFVGDSK	VVVLDEPTSG
muABCL	ERLIRDVGLT	LK.RDTQTRH	LSGGMQRKLS	VAIAFVGGSR	VVIMDEPTAG
	1101				1150
muABC1	VDPYSRRGIW	ELLLKYRQGR	TIILSTHHMD	EADILGDRIA	IISHGKLCV
muABCR	VDPYSRRSIW	DLLLKYRSGR	TIIMSTHHMD	EADLLGDRIA	IISQGRLYCS
muABCL	VDPASRRGIW	ELLLKYREGR	TLILSTHHLD	EAELLGDRVA	MVAGGSLCCC
	1151				1200
muABC1	GSSLFLKNQL	GTGYLTLVK	KDVESLSSC	RNSSSTVSCL	KKEDSVSQSS
muABCR	GTPLFLKNCF	GTGFYTLVR	KMKNIQSQRG	.GCEGVCSC	SKGFSTR...
muABCL	GSPLFLRRHL	GCGYLTLVK	SSQSLVTHDA	KGDSEDPRE	KKSDGNRTS
	1201				1250
muABC1	.....S	DAGLGSDHES	DTLTIDVSAI	SNLIRKHVSE	ARLVEDIGHE
muABCR	.....C	PTRVDEITEE	QVLDGDVQEL	MDLVYHHVPE	AKLVECIGQE
muABCL	DTAFTRGTS	KSNQAPAPGA	VPITPSTARI	LELVQQHVP	AQLVEDLPHE
	1251				1300
muABC1	LTYPVPEAA	KEGAFVELFH	EIDRLSDLG	ISSYGISETT	LEEIFLKVAE
muABCR	LIFLLPNKNF	KORAYASLFR	ELEETLADLG	LSSFGISDTP	LEEIFLKVTE
muABCL	LLLVLPHYAG	LDGSFAMVFQ	ELDQQLLELL	LTGYGISDTN	LEEIFLKVVE
	1301				1350
muABC1	ESGVDAETSD	GTLPARNRNR	AFGDKQSCLH	PFTEDDAVDP	NDSIDDPESR
muABCR	DAGAGSMFVG	GAQQKREQAG	LRHPCSAPE	KLRQYAQAPH	TCSPGQVDP
muABCL	DA...HREGG	DSRPQLHLR.	.TCTPQPPTG	PEASVLENGE	LAK...LV..L
	1351				1400
muABC1	ETDLLSGMDG	KGSYQLKGWK	LTQQQFVALL	WKRLLIARRS	RKGFFAQIVL
muABCR	KGQPSPEPED	PGVPFNTGAR	LILQHVQALL	VKRFHHTIRS	RKDFVAQIVL
muABCL	D.PQAPQGLA	PNAAQVQGW	LTCQQLRALL	HKRFLARRS	RRGLFAQVVL
	1401				1450
muABC1	PAVFVCIALV	FSLIVPPFGK	YPSLELQPWM	YNEQYTFVSN	DAPEDMGTQE
muABCR	PATFVFLALM	LSIIVPPFGE	FPALTLPWM	YGHQYTFFSM	DEPNNEHLEV
muABCL	PALFVGLALF	FSLIVPPFGQ	YPPLQLSPAM	YGPQVSFFSE	DAPGDPNRMK
	1451				1500
muABC1	LLNALTKDPG	FGTRCMEGNP	IPDTPCLAG.	EEDWTISPVP	QSIVDLFQNG
muABCR	LADVLLNRPG	FGNRCLKBEW	LPEYPCIN..	ATSWKTSPVS	PNITHLFQKQ
muABCL	LLEALLGEAG	LQEPSMQDKD	ARGSECTHSL	ACYFTVPEVP	PDVASILASG

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIG. 4D

	1501				1550
muABC1	NWTMKNPSPA	CQCSSDKIKK	MLPVCPPGAG	GLPPPQRKQK	TADILQNLTG
muABCR	KWTAAHSPSP	CKCSTREKLT	MLPECPEGAG	GLPPPQRTQR	STEVLQDLTN
muABCL	NWTPESPSPA	CQCSQPGARR	LLPDCPAGAG	GPPPPQAVAG	LGEVVQNLTG
	1551				1600
muABC1	RNISDYL VKT	YVQIIAKSLK	NKIWVNEFRY	GGFSLGVSNS	QALPPSHEVN
muABCR	RNISDYL VKT	YPALIRSSLK	SKFWVNEQRY	GGISIGGKLP	AIPISGEALV
muABCL	RNVSDFL VKT	YPSLVRRGLK	TKKWVDEVRY	GGFSLGGRDP	.DLPTGHEVV
	1601				1650
muABC1	DAIKQMKKLL	KLTKDTSADR	FLSSLGRFMA	GLDTKNNVKV	WFNNKGWHAI
muABCR	GFLSGLGQMM	NVSGGPVTRE	ASKEMLDFLK	HLETTDNIKV	WFNNKGWHAL
muABCL	RTLAEIRALL	SPQPGNALDR	ILNNLTQWAL	GLDARNSLKI	WFNNKGWHAM
	1651				1700
muABC1	SSFLNVINNA	ILRANLQKGE	NPSQYGITAF	NHPLNLTQQ	LSEVALMTTS
muABCR	VSFLNVAHNA	ILRASLPRDR	DPEEYGITVI	SQPLNLTKEQ	LSDITVLTTT
muABCL	VAFVNRRANG	LLHALLPSGP	VRHAHSITTL	NHPLNLTKEQ	LSEATLIASS
	1701				1750
muABC1	VDVLVSICVI	FAMSFVPASF	VVFLIQERVS	KAKHLQFISG	VKPVIIWLSN
muABCR	VDAVVAICVI	FAMSFVPASF	VLYLIQERT	KAKHLQFISG	VSPTTYWLTN
muABCL	VDVLVSICVV	FAMSFVPASF	TLVLIEERIT	RAKHLQLVSG	LPQTLYLWLN
	1751				1800
muABC1	FVWDMCNYVV	PATLVIIIFI	CFQQKSYVSS	TNLPVLALLL	LLYGWSITPL
muABCR	FLWDMCNYAV	SAGLVVGIFI	GFQKKAYTSP	DNLPALVSLL	MLYGWAVIPM
muABCL	FLWDMCNYLV	AVCIVVFIFL	AFQQRAYVAP	ENLPALLLLL	LLYGWSITPL
	1801				1850
muABC1	MYPASFVFKI	PSTAYVVLTS	VNLFIGINGS	VATFVLELFT	NN.KLNDIND
muABCR	MYPASFLFEV	PSTAYVALSC	ANLFIGINSS	AITFVLELFE	NNRTLLRFNA
muABCL	MYPASFFFSV	PSTAYVVLTC	INLFIGINSS	MATFVLELLS	DQ.NLQEVSR
	1851				1900
muABC1	ILKSVFLIFP	HFCLGRGLID	MVKNQAMADA	LERFGENRFV	SPLSWDLVGR
muABCR	MLRKLLIVFP	HFCLGRGLID	LALSQAVTDV	YAQFGEEYSA	NPFQWDLIGK
muABCL	ILKQVFLIFP	HFCLGRGLID	MVRNQAMADA	FERLGDKQFQ	SPLRWDIIGK
	1901				1950
muABC1	NLFAMAVEGV	VFFLITVLIQ	YRFFIRPRPV	KAKLPPLNDE	DEDVRRERQR
muABCR	NLVAMAIEGV	VYFLLTLLIQ	HHFFLTRWIA	EPAREPVFDE	DDDVAERQR
muABCL	NLLAMMAQGP	LFLITLTLQ	HRNRLLPQSK	PRLLPPLGEE	DEDVAQERER
	1951				2000
muABC1	ILDGGGQNDI	LEIKELTKIY	RRKRKPAVDR	ICIGIPPGEK	FGLLGVNAG
muABCR	VMSGGNKTDI	LKLNELTKVY	SGSSSPAVDR	LCVGVPRGEC	FGLLGVNAG
muABCL	VTGATQGDV	LVLRLDTKVY	RGQRNPAVDR	LCLGIPPGEK	FGLLGVNAG

FIG. 4E

	2001		2050
muABC1	KSTTFKMLTG	DTPVTRGDAF	LNKNSILSNI HEVHQNMGYC PQFDAITELL
muABCR	KTTTFKMLTG	DTTVTSGDAT	VAGKSILTSI SDVHQNMGYC PQFDAIDDL
muABCL	KTSTFRMVTG	DTLPSSGEAV	LAGHNVAQER SAAHRSMGYC PQSDAIFDLL
	2051		2100
muABC1	TGREHVEFFA	LLRGVPEKEV	GKFGEWAIK LGLVKYGEKY ASNYSGGNKR
muABCR	TGREHLYLYA	RLRGVPSKEI	EKVANWGIQS LGLSLYADRL AGTYSGGNKR
muABCL	TGREHLELFA	RLRGVPEAQV	AQTALSGLVR LGLPSYADRP AGTYSGGNKR
	2101		2150
muABC1	KLSTAMALIG	GPPVVFLDEP	TTGMDPKARR FLWNCALSIV KEGRSVVLTS
muABCR	KLSTAIALTG	CPPLLLLDEP	TTGMDPQARR MLWNTIVSII REGRAVVLTS
muABCL	KLATALALVG	DPAVVFLDEP	TTGMDPSARR FLWNSLLSVV REGRSVVLTS
	2151		2200
muABC1	HSMEECEALC	TRMAIMVNGR	FRCLGSVQHL KNRFGDGYTI VVRIAGS...
muABCR	HSMEECEALC	TRLAIMVKGT	FQCLGTIQHL KYKFGDGYIV TMKIKSPKDD
muABCL	HSMEECEALC	TRLAIMVNGR	FRCLGSSQHL KGRFGAGHTL TLRVPPD...
	2201		2250
muABC1	.NPDLKPVQE	FFGLAFPGSV	LKEKHRNMLQ YQLPSS.LSS LARIFSILSQ
muABCR	LLPDLNPVEQ	FFQGNFPGSV	QRERHHSMLQ FQVPS...SS LARIFQLLIS
muABCL	.Q..PEPAIA	FIRITFPGAE	LREVHGSRLR FQLPPGGRCT LTRVFRELAA
	2251		2300
muABC1	SKKRLHIEDY	SVSQTTLQDV	FVNFAKDQSD DDHLKDLSLH KNQTVVDVAV
muABCR	HKDSLLEIYY	SVTQTTLDQV	FVNFAKQOTE TYDLPLHPRA AGASWQAKLE
muABCL	QGRAHGVEDF	SVSQTTLLEV	FLYFSKDQGE EEESSRQEAE EEVSKPGRQ
	2301		2337
muABC1	LTSFLQDEKV	KESYV.....	.....
muABCR	EKSGRLQTQE	PLPAGSEQLA	NGSNPTAAED KHTRSPQ
muABCL	HPKRVSRLFLE	DPSSVETMI.	.....

FIG. 5A

	1				50
huABC1	MACWPQLRLL	LWKNLTFRRR	QTCQLLLEVA	WPLFIFLILI	SVRLSYPPYE
huABCR	MGFVRQIQLL	LWKNWTLRKR	QKIRFVVELV	WPLSLFLVLI	WLRNANPLYL
huABCL	MAFWTQLMLL	LWKNFMYRRR	QPVQLLVELL	WPLFLFFILV	AVRHSHPPLE
	51				100
huABC1	QHECHFPNKA	MPSAGTLPWV	QGIICNANNP	CFRYPTPGEA	PGVVGNFNKS
huABCR	HHECHFPNKA	MPSAGMLPWL	QGIFCNVNNP	CFQSPTPGES	PGIVSNYNNS
huABCL	HHECHFPNKP	LPSAGTVPWL	QGLICNVNNT	CFPQLTPGEE	PGRLSNFNDS
	101				150
huABC1	IVARLFSDAR	RLLLYSQKDT	SMKDMRKVLR	TLQQIKKSS.	.....SN
huABCR	ILARVYRDFQ	ELLMNAPESQ	HLGRIWTELH	ILSQFMDTLR	THPERIAGRG
huABCL	LVSRLLDAR	TVLGGASAGR	TLAGLGKLI	TLRAARSTA.	.....
	151				200
huABC1	LKLQDFLVDN	ETFSGFLYHN	LSLPKSTVVK	MLRADVILHK	VFLQGYQLHL
huABCR	IRIRDILKDE	ETLTLFLIKN	IGLSDSVVYL	LINSQVRPEQ	FAHGVPDLAL
huABCL	.....	.....	.....	.....	.....
	201				250
huABC1	TSLCNGSKSE	EMIQLG....	..DQEVSELC	GLPREKLAAA	ERVLRSNMDI
huABCR	KDIACSEALL	ERFIIFSQRR	GAKTVRYALC	SLSQGTQLWI	EDTLYANVDF
huABCL	.....	.....	.....	QPQPTKQSPL	EPPMLDVAEL
	251				300
huABC1	LKPILRTLNS	TSPFPSKELA	EATKTLLHSL	GTLAQELFSM	RSWSDMRQEV
huABCR	FKLFRVLPTL	LDQRSQGINL	RSWGGILSDM	SPRIQEFIHR	PSMQDLLWVT
huABCL	LTSLLRTESE	GLALGQA..Q	EPLHSLLEAA	EDLAQELLAL	RSLVELRALL
	301				350
huABC1	MFLTNVNSSS	SSTQIYQAVS	RIVCGHPEGG	GLKIKSLNWX	EDNNYKALFG
huABCR	RPLMQNGGPE	TFTKLMGILS	DLICGYPEGG	GSRVLSFNWX	EDNNYKALFG
huABCL	QRPRGTS GP.	.....LELLS	EALCSVRGPS	STVGPSLNWX	EASDLMELVG
	351				400
huABC1	GNGTEEDAET	FYDNSTTPYC	NDLMKNLESS	PLSRIIWKAL	KPLLVGKILY
huABCR	IDSTRKDPIY	SYDRRTTSFC	NALIQSLESN	PLTKIAWRAA	KPLLMGKILY
huABCL	....QEPESA	LPDSSLSPAC	SELIGALDSH	PLSRLLWRR	KPLILGKLLF
	401				450
huABC1	TPDTPATRQV	MAEVNKTQFE	LAVFHDLEGM	WEELSPKIWT	FMENSQEMDL
huABCR	TPDSPAARRI	LKNANSTFEE	LEHVRKLVKA	WEEVGPQIWX	FFDNSTQNMN
huABCL	APDTPFTRKL	MAQVNRTFEE	LTLLRDVREV	WEMLGPRIFT	FMNDSSNVAM
	451				500
huABC1	VRMLLDSDRN	DHFWEQQLDG	LDWTAQDIVA	FLAKHPEDVQ	SSNGSVYTWR
huABCR	IRDTLGNPTV	KDFLNRQLGE	EGITAEAILN	FLYKGPRESQ	ADDMANFDWR
huABCL	LQRLQLMQDE	GRRQPRPGGR	D.....	.HMEALRSFL	DPGSGGYSWQ

FIG. 5A

FIG. 5B

	501		550
huABC1	BAFNETNQAI	RTISRFMECV	NLNKLEPIAT EVWLINKSME LLDERKFWAG
huABCR	DIFNITDRTL	RLVNQYLECL	VLDKFESYND ETQLTQRALS LLEENMFWAG
huABCL	DAHADVGHV	GTGGRVTECL	SLDKLEAAPS EAALVSRALQ LLAEHRFWAG
	551		600
huABC1	IVFTG.....	...ITPGSIE	LPHHVKYKIR MDIDNVERTN KIKDGYWDPG
huABCR	VVFPD.....	...MYPWTSS	LPPHVKYKIR MDIDVVEKTN KIKDRYWDG
huABCL	VVFLGPDSS	DPTEHPTPDL	GPGHVRIKIR MDIDVVTRTN KIRDRFWDPG
	601		650
huABC1	PRADPFEDMR	YVWGGFAYLQ	DVVEQAIIRV LTGTEKKTGV YMQQMPYPCY
huABCR	PRADPVEDFR	YIWGGFAYLQ	DMVEQGITS QVQAEAPVGI YLQQMPYPCF
huABCL	PAADPLTDLR	YVWGGFVYLQ	DLVERAAVRV LSGANPRAGL YLQQMPYPCY
	651		700
huABC1	VDDIFLRVMS	RSMPLFMTLA	WIYSVAVIK GIVYEKEARL KETMRIMGLD
huABCR	VDDSFMIILN	RCFPIMVLA	WIYSVSMTVK SIVLEKELRL KETLKNQGS
huABCL	VDDVFLRVLS	RSLPLFLTTL	WIYSVTLTVK AVVREKETRL RDTMRAMGLS
	701		750
huABC1	NSILWFSWFI	SSLIPLLVSA	GLLVVILKLG NLLPYSDPSV VFVFLSVFAV
huABCR	NAVIWCTWFL	DSFSIMSMI	FLLTIFIMHG RILHYSDPFI LFLFLLAFST
huABCL	RAVLWLWFL	SCLGPFLLSA	ALLVLVLKLG DILPYSHPGV VFLFLAAFAV
	751		800
huABC1	VTILQCFLIS	TLFSRANLAA	ACGGIIYFTL YLPYVLCVAW QDYVGFTLKI
huABCR	ATIMLCFLLS	TFFSKASLAA	ACSGVIYFTL YLPHILCFW QDRMTAELKK
huABCL	ATVTQSFLLS	AFFSRANLAA	ACGGLAYFSL YLPYVLCVAW RDRLPAGGRV
	801		850
huABC1	FASLLSPVAF	GFGCEYFALF	EEQGIGVQWD NLFESPVEED GFNLTTSVSM
huABCR	AVSLLSPVAF	GFGTEYLVRF	EEQGLGLQWS NIGNSPTEGD EFSFLLSMQM
huABCL	AASLLSPVAF	GFGCESLALL	EEQGEQAQWH NVG.TRPTAD VFSLAQVSGL
	851		900
huABC1	MLFDTFLYGV	MTWYIEAVFP	GQYGIPRPWY FPCTKSYWFG .....
huABCR	MLLDAAVYGL	LAWYLDQVFP	GDYGTPLPWY FLLQESYWLS GEGCSTREER
huABCL	LLLDAAALYGL	ATWYLEAVCP	GQYGIPEPWN PFRRSYWCG .....
	901		950
huABC1	.....EESDE	KSHPGSNQKR	ISEICMEEEP THLKLGVSIQ NLVKVYRDGM
huABCR	ALEKTEPLTE	ETEDPEHPEG	IHDSFFEREH PGWVPGVCVK NLVKIFEPCG
huABCL	.....PRPP	KSPAPCPTPL	DPKVLVEEAP PGLSPGVSVR SLEKRFPGPS
	951		1000
huABC1	KVAVDGLALN	FYEGQITSFL	GHNGAGKTTT MSILTGLFPP TSGTAYILGK
huABCR	RPAVDRLNIT	FYENQITAF	GHNGAGKTTT LSILTGLLPP TSGTVLVGGR
huABCL	QPALRGLSLD	FYQGHITAF	GHNGAGKTTT LSILSGLFPP SGGSAFILGH

FIG. 5C

	1001				1050
huABC1	DIRSEMSTIR	QNLGVCPQHN	VLFDMLTVEE	HIWFYARLKG	LSEKHVKAEM
huABCR	DIETSLDAVR	QSLGMCPQHN	ILFHHLTVAE	HMLFYAQLKG	KSQEEAQLEM
huABCL	DVRSSMAAIR	PHLGVCPQYN	VLFDMLTVDE	HVWFYGRLLKG	LSAAVVGPEQ
	1051				1100
huABC1	EQMALDVGLP	SSKLKSKTSQ	LSGGMQRKLS	VALAFVGGSK	VVILDEPTAG
huABCR	EAMLEDTGLH	HK.RNEEAQD	LSGGMQRKLS	VAIAFVGDAK	VVILDEPTSG
huABCL	DRLLQDVGLV	SK.QSVQTRH	LSGGMQRKLS	VAIAFVGGSQ	VVILDEPTAG
	1101				1150
huABC1	VDPYSRRGIW	ELLLKYRQGR	TIILSTHHMD	EADVLGDRIA	IISHGKLCV
huABCR	VDPYSRRSIW	DLLLKYRSGR	TIIMSTHHMD	EADLLGDRIA	IIAQGRLYCS
huABCL	VDPASRRGIW	ELLLKYREGR	TLILSTHHLD	EAELLGDRVA	VVAGGRLLCC
	1151				1200
huABC1	GSSLFLKNQL	GTGYLTLVK	KDVESLSSC	RNSSSTVSYL	KKEDSVSQSS
huABCR	GTPLFLKNCF	GTGLYTLVR	KMKNIQSQRK	...GSEGT	SCSSKGFSTT
huABCL	GSPLFLRRHL	GSGYTLVK	ARLPLTNEK	.....	..ADTMEGS
	1201				1250
huABC1	SDAGLGSDHE	SDTLTIDVSA	ISNLIRKHVS	EARLVEDIGH	ELTYVLPYEA
huABCR	CPAHVDDLTP	EQVLDGDVNE	LMDVVLHHVP	EAKLVECIGQ	ELIFLLPNKN
huABCL	VDTRQEKKNG	SQGSRVGTPQ	LLALVQHWVP	GARLVEELPH	ELVLVLPYTG
	1251				1300
huABC1	AKEGAFVELF	HEIDRLSDL	GISSYGISET	TLEEIFLKVA	EESGVDAETS
huABCR	FKHRAYASLF	RELEETLADL	GLSSFGISDT	PLEEIFLKVT	EDSDSGPLFA
huABCL	AHDGSFATLF	RELDTRLAEL	RLTGYGISDT	SLEEIFLKVV	EECAADTME
	1301				1350
huABC1	DGTLPARRRNR	RAFGDKQSCL	RPFTEDDAAD	PNDSIDPES	RETDLLSGMD
huABCR	GGAQQKREN.	..VNPRHPCL	GPREKAGQTP	QDSNVCPGA	PAAHPEGQPP
huABCL	DGSCGQHLCT	.GIAGLDVTL	RLKMPPQETA	LENCEPAGSA	PETDQSGSPD
	1351				1400
huABC1	GK...GSYQ	VKGWKLTOQQ	FVALLWKRL	IARRSRKGFF	AQIVLPAVFF
huABCR	PEPECPPGQL	NTGTQLVLQH	VQALLVKRFQ	HTIRSHKDFL	AQIVLPATFV
huABCL	AV...G..R	VQGWALTRQQ	LQALLLKRFL	LARRSRGLF	AQIVLPALFV
	1401				1450
huABC1	CIALVFSILV	PPFGKYPSE	LQPMWYNEQY	TFVSNDAPE	TGTLELLNAL
huABCR	FLALMLSIVI	LPFGYEPALT	LHPWIYGQY	TFFSMDEPGS	EQFTVLADVL
huABCL	GLALVFSILV	PPFGHYPALR	LSPTMYGAQV	SFFSEDAPGD	PGRARLLEAL
	1451				1500
huABC1	TKDPGFGTRC	MEGNPIPDTP	CQAGEEEWTT	APVPQTIMDL	FQNGNWTMQN
huABCR	LNKPGFGNRC	LKEGWLPEYP	CGN.STPWKT	PSVSPNITQL	FQKQKWTQVN
huABCL	L.....	.QEAGLEEPP	VQHSSHRFSA	PEVPAEVAKV	LASGNWTPES

bioRxiv preprint doi: <https://doi.org/10.1101/2019.04.11.343434>; this version posted April 11, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIG. 5D

	1501		1550
huABC1	PSPACQCSSD	KIKKMLPVCP	PGAGGLPPPQ RKQNTADILQ DLTGRNISDY
huABCR	PSPSCRCSTR	EKLTMLEPCP	EGAGGLPPPQ RTQRSTEILQ DLTDRNISDF
huABCL	PSPACQCSQP	GARRLLPDCP	AAAGGPPPPQ AVTGSGEVVQ NLTGRNLSDF
	1551		1600
huABC1	LVKTYVQIIA	KSLKNKIWVN	EFYGGFSLG VSNTQALPPS QEVNDATKQM
huABCR	LVKTYPALIR	SSLKSKFWVN	EQRYGGISIG GKLPVVPITG EALVGFLSDL
huABCL	LVKTYPRLVR	QGLKTKKWVN	EVRYGGFSLG GRDPGLP.SG QELGRSVEEL
	1601		1650
huABC1	KKHLKLAKDS	SADRFLNSLG	RFMTGLDTRN NVKVWFNNKG WHAISSFLNV
huABCR	GRIMNVSGGP	ITREASKEIP	DFLKHLETED NIKVWFNNKG WHALVSFLNV
huABCL	WALLSPLPGG	ALDRVLKNLT	AWAHS LDAQD SLKIWFNNKG WHSMVAFVNR
	1651		1700
huABC1	INNAILRANL	QKGENPSHYG	ITAFNHPLNL TKQQLSEVAP MTTSDVDLVS
huABCR	AHNAILRASL	PKDRSPEEYG	ITVISQPLNL TKEQLSEITV LTTSVDAVVA
huABCL	ASNAILRAHL	PPGPARHAHS	ITTLNHPLNL TKEQLSEAL MASSVDLVS
	1701		1750
huABC1	ICVIFAMSFV	PASFVVFLIQ	ERVSKAKHLQ FISGVKPIY WLSNFWDMC
huABCR	ICVIFSMSFV	PASFVLYLIQ	ERVNKS KHLQ FISGVSPPTY WVTNFLWDIM
huABCL	ICVVFAMSFV	PASFTLV LIE	ERVTRAKHLQ LMGGLSPTLY WLGNFLWDMC
	1751		1800
huABC1	NYVVPATLVI	IIFICFQOKS	YVSSTNLPVL ALLLLLYGWS ITPLMYPASF
huABCR	NYSVSAGLVV	GIFIGFQKKA	YTSPENLPAL VALLLLYGWA VIPMMYPASF
huABCL	NYLVPACIVV	LIFLAFQORA	YVAPANLPAL LLLLLLYGWS ITPLMYPASF
	1801		1850
huABC1	VFKIPSTAYV	VLTSVNLFIG	INGSVATFVL ELFTDN.KLN NINDILKSVF
huABCR	LFDVPSTAYV	ALSCANLFIG	INSSAITFIL ELFDNNRTLL RFNAVL RKLL
huABCL	FFSVPSTAYV	VLTCINLFIG	INGSMATFVL ELFS DQ.KLQ EVSRILKQVF
	1851		1900
huABC1	LIFPHFCLGR	GLIDMVKNQA	MADALERFGE NRFVSPLSWD LVGRNLFAMA
huABCR	IVFPHFCLGR	GLIDLALSQA	VTDVYARFGE EHSANPFHWD LIGKNLFAMV
huABCL	LIFPHFCLGR	GLIDMVRNQA	MADAFERLGD RQFQSPLRWE VVGKNLLAMV
	1901		1950
huABC1	VEGVVFFLIT	VLIQYRFFIR	PRPVNAKLSP LNDEDEDVRR ERQRILDGGG
huABCR	VEGVVYFLLT	LLVQRHFFLS	QWIAEPTKEP IVDEDDDAE ERQRIITGGN
huABCL	IQGPLFLLFT	LLLQHRSQLL	PQPRVRS LPL LGEED EDVAR ERERVVQGAT
	1951		2000
huABC1	QNDILEIKEL	TKIYRRKRKP	AVDRICVGIP PGE CFGLLG V NGAGKSSTFK
huABCR	KTDILRLHEL	TKIYLG TSSP	AVDRLCVGVR PGE CFGLLG V NGAGKTTTFK
huABCL	QGDVLVLRNL	TKVYRGQRMP	AVDRLC LGIP PGE CFGLLG V NGAGKTSTFR

bioRxiv preprint doi: <https://doi.org/10.1101/2020.03.10.332000>; this version posted March 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

# FIG. 5E

	2001				2050
huABC1	MLTGDTTVTR	GDAFLNRNSI	LSNIHEVHQN	MGYCPQFDAI	TELLTGREHV
huABCR	MLTGDTTVTS	GDA TVAGKSI	LTNISEVHQN	MGYCPQFDAI	DELLTGREHL
huABCL	MVTGDTLASR	GEAVLAGHSV	AREPSAAHLS	MGYCPQSDAI	FELLTGREHL
	2051				2100
huABC1	EFFALLRGVP	EKEVGKVGEW	AIRKLGLVKY	GEKYAGNYSG	GNKRKLSTAM
huABCR	YLYARLRGVP	AEIEKVANW	SIKSLGLTVY	ADCLAGTYSG	GNKRKLSTAI
huABCL	ELLARLRGVP	EAQVAQTAGS	GLARLGLSWY	ADRPAGTYSG	GNKRKLATAL
	2101				2150
huABC1	ALIGGPPVVF	LDEPTTGMDP	KARRFLWNCA	LSVVKEGRSV	VLTSHSMEEC
huABCR	ALIGCPPLVL	LDEPTTGMDP	QARRMLWNV I	VSIIREGRAV	VLTSHSMEEC
huABCL	ALVGDPVAVF	LDEPTTGMDP	SARRFLWNSL	LAVVREGRSV	MLTSHSMEEC
	2151				2200
huABC1	EALCTRMAIM	VNGRFRCLGS	VQHLKNRFGD	GYTIVVRIAG	S...NPDLK
huABCR	EALCTRLAIM	VKGAFRCMGT	IQHLKSKFGD	GYIVTMKIKS	PKDDL LPDLN
huABCL	EALCSRLAIM	VNGRFRCLGS	PQHLKGRFAA	GHTLT LRVPA	A.....RSQ
	2201				2250
huABC1	PVQDFFGLAF	PGSVPKEKHR	NMLQYQLPSS	.LSSLARIFS	ILSQSKKRLH
huABCR	PVEQFFQGNF	PGSVQRRHY	NMLQFQVSS.	..SSLARIFQ	LLLSHKDSLL
huABCL	PAAAFVAAEF	PGSELREAHG	GRLRFQLPPG	GRCALARVFG	ELAVHGAEHG
	2251				2300
huABC1	IEDYSVSQTT	LDQVFVNFAK	DQSDDHLKD	LSLHKNQTVV	DVAVLTSFLQ
huABCR	IEEYSVTQTT	LDQVFVNFAK	QQTESHDLPL	HPRAAGASRQ	AQD.....
huABCL	VEDFSVSQTM	LEEVFLYFSK	DQGKDEDT EE	QKEAGVGVDP	APGLQHPKRV
	2301				
huABC1	DEKVKESYV.	....			
huABCR	.....	....			
huABCL	SQFLDDPSTA	ETVL			